Routh's procedure for non-Abelian symmetry groups

M. Crampin[1] and T. Mestdag[1,2,a)]
[1]*Department of Mathematical Physics and Astronomy, Ghent University, Krijgslaan 281, B-9000 Ghent, Belgium*
[2]*Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, Michigan 48109, USA*

(Received 6 November 2007; accepted 1 February 2008; published online 5 March 2008)

We extend Routh's reduction procedure to an arbitrary Lagrangian system (that is, one whose Lagrangian is not necessarily the difference of kinetic and potential energies) with a symmetry group which is not necessarily Abelian. To do so, we analyze the restriction of the Euler–Lagrange field to a level set of momentum in velocity phase space. We present a new method of analysis based on the use of quasivelocities. We discuss the reconstruction of solutions of the full Euler–Lagrange equations from those of the reduced equations. © *2008 American Institute of Physics.* [DOI: 10.1063/1.2885077]

I. INTRODUCTION

Routh's procedure, in its original form (as described in his treatise[13]), was concerned with eliminating from a Lagrangian problem the generalized velocities corresponding to so-called ignorable or cyclic coordinates. Let L be a Lagrangian on $\mathbf{R}^n$ that does not explicitly depend on m of its base variables, say, the coordinates θ^a. From the Euler–Lagrange equations for these coordinates,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot\theta^a}\right)-\frac{\partial L}{\partial\theta^a}=0,$$

we can immediately conclude that the functions $\partial L/\partial\dot\theta^a$ are constants, say,

$$\frac{\partial L}{\partial\dot\theta^a}=\pi_a.$$

These equations express the conservation of generalized momentum. Routh's idea is to solve these equations for the variables $\dot\theta^a$ and to introduce what he calls the "modified Lagrangian function," the restriction of the function

$$L'=L-\frac{\partial L}{\partial\dot\theta^a}\dot\theta^a$$

to the level set where the momentum is π_a. One can easily verify that the $(n-m)$ Euler–Lagrange equations for the remaining variables x^i can be rewritten as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot x^i}\right)-\frac{\partial L}{\partial x^i}=0\Rightarrow\frac{d}{dt}\left(\frac{\partial L'}{\partial\dot x^i}\right)-\frac{\partial L'}{\partial x^i}=0.$$

For example, if the Lagrangian takes the form

[a)]Electronic mail: tom.mestdag@ugent.be.

$$L(x,\theta,\dot{x},\dot{\theta}) = \tfrac{1}{2}k_{ij}(x)\dot{x}^i\dot{x}^j + k_{ia}(x)\dot{x}^i\dot{\theta}^a + \tfrac{1}{2}k_{ab}(x)\dot{\theta}^a\dot{\theta}^b - V(x),$$

the conservation of momentum equations read $k_{ia}\dot{x}^i + k_{ab}\dot{\theta}^b = \pi_a$, and they can be solved for the variables $\dot{\theta}^a$ if (k_{ab}) is a nonsingular matrix. The modified Lagrangian function is

$$L'(x,\dot{x}) = \tfrac{1}{2}(k_{ij} - k^{ab}k_{ia}k_{jb})\dot{x}^i\dot{x}^j + k^{ab}k_{ia}\pi_b\dot{x}^i - \left(V + \tfrac{1}{2}k^{ab}\pi_a\pi_b\right),$$

where k^{ab} denotes a component of the matrix inverse to (k_{ab}), in the usual way. Clearly, the advantage of this technique is that the reduced equations in L' involve only the unknowns x^i and $\dot{x}^i$; they can, in principle, be directly solved for the x^i, and the θ^a may then be found (if required) from the momentum equation.

A modern geometric interpretation of this reduction procedure can be found in, e.g., Ref. 8. The above Lagrangian L is of the form $T-V$, where the kinetic energy part is derived from a Riemannian metric (i.e., we are dealing with a so-called simple mechanical system). The function L is defined on the tangent manifold of a manifold of the form $M = S \times G$ (in this case $\mathbf{R}^n$) and it is invariant under an Abelian Lie group G (in this case the group of translations $\mathbf{R}^m$). The main feature of the procedure is that the modified Lagrangian function and its equations can be defined in terms of the coordinates on S only. However, to give the definition of the modified function an intrinsic meaning, we should define this function, from now on called the Routhian, rather as the restriction to a level set of momentum of

$$\mathcal{R} = L - \frac{\partial L}{\partial \dot{\theta}^a}(\dot{\theta}^a + \Lambda^a_i \dot{x}^i),$$

with $\Lambda^a_i = k^{ab}k_{ib}$, i.e.,

$$\mathcal{R}(x,\dot{x}) = \tfrac{1}{2}(k_{ij} - k^{ab}k_{ia}k_{jb})\dot{x}^i\dot{x}^j - \left(V + \tfrac{1}{2}k^{ab}\pi_a\pi_b\right).$$

The coefficients Λ^a_i form a connection on the trivial principal bundle $M = S \times G \to S$, usually called the mechanical connection, and $\dot{\theta}^a + \Lambda^a_i \dot{x}^i$ is, in fact, the vertical projection of the vector $(\dot{x}^i, \dot{\theta}^a)$. The $(n-m)$ Euler–Lagrange equations in x^i then become

$$\frac{d}{dt}\left(\frac{\partial \mathcal{R}}{\partial \dot{x}^i}\right) - \frac{\partial \mathcal{R}}{\partial x^i} = -B^a_{ij}\pi_a\dot{x}^j,$$

where in the term on the right-hand side

$$B^a_{ij} = \frac{\partial \Lambda^a_i}{\partial x^j} - \frac{\partial \Lambda^a_j}{\partial x^i}$$

has a coordinate-free interpretation as the curvature of the connection.

In Refs. 9 and 10, Marsden *et al.* extended the above procedure to the case of simple mechanical systems with a non-Abelian symmetry group G and where the base manifold has a principal bundle structure $M \to M/G$. The procedure has recently been further extended to cover Lagrangian systems in general by Castrillon-Lopez.[1]

The most important contribution of our paper lies in the geometric formalism we will adopt. The bulk of the literature dealing with different types of reduction of Lagrangian systems has relied heavily on methods coming from the calculus of variations. In fact, as in, e.g., Refs. 1, 5, and 10, the reduced equations of motion are usually obtained by considering some reduced version of Hamilton's principle. Our method is different from those of other authors in that it does not involve consideration of variations. It is distinctively Lagrangian (as opposed to Hamiltonian) and is based on the geometrical analysis of regular Lagrangian systems, where solutions of the Euler–Lagrange equations are interpreted as integral curves of an associated second-order differential equation field on the velocity phase space, that is, the tangent manifold of the configuration space. Consequently, our derivation of Routh's equations is relatively straightforward and is a natural extension of that used by Routh in the classical case. In particular, we will show how Routh's equations can be derived directly from the Euler–Lagrange equations by choosing a suitable adapted frame or, equivalently, by employing well-chosen quasivelocities. This line of thinking has already provided some new insights into, e.g., the geometry of second-order differential systems with symmetry.[3]

We deal from the beginning with arbitrary Lagrangians, i.e., Lagrangians not necessarily of the form $T-V$.

As in Ref. 10, we explain how solutions of the Euler–Lagrange equations with a fixed momentum can be reconstructed from solutions of the reduced equations. The method relies on the availability of a principal connection on an appropriate principal fiber bundle. We will introduce, in fact, two connections that serve the same purpose.

We describe the basic features of our approach in Sec. II. The reduction of a Lagrangian system to a level set of momentum is discussed in Sec. III, and our generalization of Routh's procedure is explained there. Section IV contains some general remarks about using a principal connection to reconstruct an integral curve of a dynamical vector field from one of a reduction of it. In Sec. V, we describe the two principal connections that can be used in the specific reconstruction problem we are concerned with, while in Sec. VI, we carry out the reduction in detail, first in the Abelian case, then in general. In Sec. VII, we specialize to simple mechanical systems, in order to compare our results with those published elsewhere. We conclude the paper with a couple of illustrative examples.

II. PRELIMINARIES

We will be concerned with Lagrangian systems admitting non-Abelian (that is to say, not necessarily Abelian) symmetry groups. We begin by explaining what assumptions we make about the action of a symmetry group.

We will suppose that $\psi^M : G\times M \to M$ is a free and proper left action of a connected Lie group G on a manifold M. It should be noticed from the outset that this convention differs from the one in, e.g., Refs. 3 and 6, but resembles the one taken in, e.g., Refs. 8 and 10.

With such an action, M is a principal fiber bundle with group G; we write M/G for the base manifold and $\pi^M : M \to M/G$ for the projection. We denote by $\mathfrak{g}$ the Lie algebra of G. For any $\xi\in\mathfrak{g}$, $\tilde{\xi}$ will denote the corresponding fundamental vector field on M, that is, the infinitesimal generator of the one-parameter group $\psi^M_{\exp(t\xi)}$ of transformations of M. The Lie bracket of two fundamental vector fields satisfies $[\tilde{\xi},\tilde{\eta}]=-\widetilde{[\xi,\eta]}$ (see, e.g., Ref. 8). Since G is connected, a tensor field on M is invariant under the action of G if and only if its Lie derivatives by all fundamental vector fields vanish. In particular, a vector field X on M is invariant if and only if $[\tilde{\xi},X]=0$ for all $\xi\in\mathfrak{g}$. We will usually work with a fixed basis for $\mathfrak{g}$, which we denote by $\{E_a\}$; then, for X to be invariant, it is enough that $[\tilde{E}_a,X]=0$, $a=1,2,\ldots,\dim(\mathfrak{g})$.

We suppose that we have at our disposal a principal connection on M. For the most part, it will be convenient to work with connections in the following way. A connection is a left splitting of the short exact sequence

$$0 \to M\times\mathfrak{g} \to TM \to (\pi^M)^* T(M/G) \to 0$$

of vector bundles over M; we identify $M\times\mathfrak{g}$ with the vertical sub-bundle of $TM\to M$ by $(m,\xi)\mapsto \tilde{\xi}|_m$. Thus, we may think of a connection as a type (1,1) tensor field ω on M, which is a projection map on each tangent space, with the image tangent to the fiber of π^M. The connection is principal just when ω is invariant, that is, when $\mathcal{L}_{\tilde{\xi}}\omega=0$ for all $\xi\in\mathfrak{g}$. The kernel distribution of ω is the horizontal distribution of the connection. An alternative test for the invariance of the connection is that its horizontal distribution should be invariant (as a distribution); that is, for any horizontal vector field X, $[\tilde{\xi},X]$ is also horizontal for all ξ. We will often refer to a connection by the symbol of the corresponding tensor field.

Let $\{X_i\}$ be a set of local vector fields on M which are linearly independent, horizontal with respect to ω, and invariant. Such a set of vector fields consists of the horizontal lifts of a local basis of vector fields on M/G and, in particular, we may take for the X_i the horizontal lifts of coordinate fields on M/G. We then have a local basis $\{X_i, \widetilde{E}_a\}$ of vector fields on M. We will very often work with such a basis, which we call a standard basis. The Lie brackets of pairs of vector fields in a standard basis are

$$[X_i,X_j] = R^a_{ij}\widetilde{E}_a, \quad [X_i,\widetilde{E}_a] = 0, \quad [\widetilde{E}_a,\widetilde{E}_b] = -C^c_{ab}\widetilde{E}_c.$$

The R^a_{ij} are the components of the curvature of ω, regarded as a $\mathfrak{g}$-valued tensor field. The second relation simply expresses the invariance of the X_i. In the third expression, the C^c_{ab} are structure constants of $\mathfrak{g}$ with respect to the chosen basis.

It will sometimes be convenient to have also a basis $\{X_i, \hat{E}_a\}$ that consists entirely of invariant vector fields. Let $U \subset M/G$ be an open set over which M is locally trivial. The projection π^M is locally given by projection onto the first factor in $U \times G \to U$, and the (left) action by $\psi^M_g(x,h) = (x,gh)$. The vector fields on M defined by

$$\hat{E}_a:(x,g) \mapsto (\widetilde{\mathrm{ad}_g E_a})(x,g) = \psi^{TM}_g(\widetilde{E}_a(x,e)).$$

(where e is the identity of G) are invariant. The relation between the sets $\{\hat{E}_a\}$ and $\{\widetilde{E}_a\}$ can be expressed as $\hat{E}_a(x,g) = \mathcal{A}^b_a(g)\widetilde{E}_b(x,g)$, where $(\mathcal{A}^b_a(g))$ is the matrix representing ad_g with respect to the basis $\{E_a\}$ of $\mathfrak{g}$. In particular, $\mathcal{A}^b_a(e) = \delta^b_a$. Since $[\widetilde{E}_a, \hat{E}_b] = 0$, the coefficients $\mathcal{A}^b_a$ have the property that $\widetilde{E}_a(\mathcal{A}^c_b) = C^c_{ad}\mathcal{A}^d_b$.

We revert to consideration of a standard basis. We define the component 1-forms ω^a of the tensor field ω by $\omega = \omega^a\widetilde{E}_a$. Then, $\omega^a(X_i) = 0$, $\omega^a(\widetilde{E}_b) = \delta^a_b$. Thus, the ω^a comprise part of the basis of 1-forms dual to the standard basis. We denote by ϑ^i the remaining 1-forms in the dual basis.

Most of the objects of interest, such as the Lagrangian and the corresponding Euler–Lagrange field Γ, live on the tangent manifold of M, which we denote by $\tau: TM \to M$. We recall that there are two canonical ways of lifting a vector field, say, Z, from M to TM. The first is the complete or tangent lift, Z^C, whose flow consists of the tangent maps of the flow of Z. The second is the vertical lift, Z^V, which is tangent to the fibers of τ and on the fiber over m coincides with the constant vector field Z_m. We have $T\tau(Z^C) = Z$, while $T\tau(Z^V) = 0$. Moreover, TM is equipped with a canonical type (1,1) tensor field called the vertical endomorphism and denoted by S, which is essentially determined by the fact that $S(Z^C) = Z^V$ and $S(Z^V) = 0$. For more details on this material, see, e.g., Refs. 4 and 16. The set $\{X_i^C, \widetilde{E}_a^C, X_i^V, \widetilde{E}_a^V\}$, consisting of the complete and vertical lifts of $\{X_i, \widetilde{E}_a\}$, forms a local basis of vector fields on TM.

Let $\{Z_\alpha\}$ be a local basis of vector fields on M, and $\{\theta^\alpha\}$ the dual basis of 1-forms. These 1-forms define fiber-linear functions $\vec{\theta}^\alpha$ on TM, such that for any $u \in T_mM$, $u = \vec{\theta}^\alpha(u)Z_\alpha(m)$. These functions are therefore the components of velocities with respect to the specified vector-field basis. We may use these functions as fiber coordinates. Coordinates of this type are sometimes called quasivelocities, and we will use this terminology. In the case of interest, we have a standard basis $\{X_i, \widetilde{E}_a\}$ and its dual $\{\vartheta^i, \omega^a\}$; we denote the corresponding quasivelocities by $v^i = \vec{\vartheta}^i$, $v^a = \vec{\omega}^a$.

We will need to evaluate the actions of the vector fields X_i^C, $\widetilde{E}_a^C$, X_i^V, and $\widetilde{E}_a^V$ on v^i and v^a. Now, for any vector field Z and 1-form θ on M,

$$Z^C(\vec{\theta}) = \overrightarrow{\mathcal{L}_Z\theta}, \quad Z^V(\vec{\theta}) = \tau^*\theta(Z).$$

Most of the required results are easy to derive from these formulas. The only tricky calculation is that of $X_i^C(\vec{\omega}^a)$, for which we need the Lie derivative of a connection form by a horizontal vector field. We have

$$(\mathcal{L}_{X_i}\omega^a)(\widetilde{E}_b) = X_i(\delta^a_b) - \omega^a([X_i,\widetilde{E}_b]) = 0,$$

$$(\mathcal{L}_{X_i}\omega^a)(X_j) = -\omega^a([X_i,X_j]) = -R^a_{ij}.$$

In the first, we have used the invariance of the horizontal vector fields. In summary, the relevant derivatives of the quasivelocities are

$$X_i^C(v^j) = 0, \quad X_i^V(v^j) = \delta_i^j, \quad X_i^C(v^a) = -R^a_{ij}v^j, \quad X_i^V(v^a) = 0,$$

$$\widetilde{E}_a^C(v^i) = 0, \quad \widetilde{E}_a^V(v^i) = 0, \quad \widetilde{E}_a^C(v^b) = C^b_{ac}v^c, \quad \widetilde{E}_a^V(v^b) = \delta_a^b.$$

Finally, we list some important Lie brackets of the basis vector fields:

$$[\widetilde{E}_a^C, X_i^C] = [\widetilde{E}_a, X_i]^C = 0, \quad [\widetilde{E}_a^C, X_i^V] = [\widetilde{E}_a, X_i]^V = 0,$$

$$[\widetilde{E}_a^C, \widetilde{E}_b^C] = [\widetilde{E}_a, \widetilde{E}_b]^C = -C^c_{ab}\widetilde{E}_c^C, \quad [\widetilde{E}_a^C, \widetilde{E}_b^V] = [\widetilde{E}_a, \widetilde{E}_b]^V = -C^c_{ab}\widetilde{E}_c^V.$$

III. THE GENERALIZED ROUTH EQUATIONS

We begin by explaining, in general terms, how we will deal with the Euler–Lagrange equations.

Consider a manifold M, with local coordinates (x^α), and its tangent bundle $\tau: TM \to M$, with corresponding local coordinates (x^α, u^α). A Lagrangian L is a function on TM; its Euler–Lagrange equations,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial u^\alpha}\right) - \frac{\partial L}{\partial x^\alpha} = 0,$$

comprise a system of second-order ordinary differential equations for the extremals; in general, the second derivatives $\ddot{x}^\alpha$ are given implicitly by these equations. We say that L is regular if its Hessian with respect to the fiber coordinates,

$$\frac{\partial^2 L}{\partial u^\alpha \partial u^\beta},$$

considered as a symmetric matrix, is everywhere nonsingular. When the Lagrangian is regular, the Euler–Lagrange equations may be solved explicitly for $\ddot{x}^\alpha$, and so determine a system of differential equations of the form $\ddot{x}^\alpha = f^\alpha(x,\dot{x})$. These equations can in turn be thought of as defining a vector field Γ on TM, a second-order differential equation field, namely,

$$\Gamma = u^\alpha \frac{\partial}{\partial x^\alpha} + f^\alpha \frac{\partial}{\partial u^\alpha}.$$

We call this the Euler–Lagrange field of L. The Euler–Lagrange equations may be written as

$$\Gamma\left(\frac{\partial L}{\partial u^\alpha}\right) - \frac{\partial L}{\partial x^\alpha} = 0,$$

and when L is regular, these equations, together with the assumption that it is a second-order differential equation field, determine Γ.

This is essentially how we will deal with the Euler–Lagrange equations throughout: that is, we will assume that L is regular and we will work with the Euler–Lagrange field Γ and with the Euler–Lagrange equations in the form given above. However, we need to be able to express those equations in terms of a basis of vector fields on M which is not necessarily of coordinate type. It is easy to see that if $\{Z_\alpha\}$ is such a basis, then the equations,

$$\Gamma(Z_\alpha^V(L)) - Z_\alpha^C(L) = 0,$$

are equivalent to the Euler–Lagrange equations. The fact that Γ is a second-order differential equation field means that it takes the form

$$\Gamma = w^\alpha Z_\alpha^C + \Gamma^\alpha Z_\alpha^V,$$

where w^α are the quasivelocities corresponding to the basis $\{Z_\alpha\}$.

We now build in the assumption that L has a symmetry group G, which acts in such a way that M is a principal bundle with G as its group, as we described above. We will suppose that the Lagrangian is invariant under the induced action of G on TM. This tangent action is defined by the collection of transformations $\psi_g^{TM} = T\psi_g^M$ on TM, $g \in G$. By construction, the fundamental vector fields for this induced action are the complete lifts of the fundamental vector fields of the action on M; the invariance of the Lagrangian can therefore be characterized by the property $\tilde{E}_a^C(L)=0$. We have shown in Ref. 11 that if L is invariant, then so also is Γ, which is to say that $[\tilde{E}_a^C, \Gamma]=0$.

We choose a principal connection on M and a basis of vector fields $\{X_i, \tilde{E}_a\}$ adapted to it (a standard basis), as described above. Then, the Euler–Lagrange equations for L are

$$\Gamma(X_i^V(L)) - X_i^C(L) = 0,$$

$$\Gamma(\tilde{E}_a^V(L)) - \tilde{E}_a^C(L) = 0.$$

However, by assumption, $\tilde{E}_a^C(L)=0$: it follows immediately that $\Gamma(\tilde{E}_a^V(L))=0$. So, the functions $\tilde{E}_a^V(L)$ are first integrals, which clearly generalize the momenta conjugate to ignorable coordinates in the classical Routhian picture. We write p_a for $\tilde{E}_a^V(L)$. The Euler–Lagrange field is tangent to any submanifold $p_a=\mu_a=\text{const}$, $a=1,2,\ldots,\dim(\mathfrak{g})$, that is, any level set of momentum. By a well-known argument (see, e.g., Ref. 8), we may regard $(x,v)\mapsto(p_a(x,v))$ as a map from TM to $\mathfrak{g}^*$, the dual of the Lie algebra $\mathfrak{g}$, and this map is equivariant between the given action of G on TM and the coadjoint action of G on $\mathfrak{g}^*$ (the coadjoint action is defined as $\langle\xi, \mathrm{ad}_g^*\,\mu\rangle = \langle \mathrm{ad}_g\,\xi, \mu\rangle$). We have

$$\tilde{E}_a^C(p_b) = \tilde{E}_a^C\tilde{E}_b^V(L) = [\tilde{E}_a^C, \tilde{E}_b^V](L) = -C_{ab}^c\tilde{E}_c^V(L) = -C_{ab}^c p_c,$$

which expresses this result in our formalism.

We will also need a less coordinate-dependent version of the Hessian. In fact, the Hessian of L at $w\in TM$ is the symmetric bilinear form g on T_mM, $m=\tau(w)$, given by $g(u,v)=u^Vv^V(L)$, where the vertical lifts are to w. We can equally well regard g as a bilinear form on the vertical subspace of T_wTM, by identifying u and v with their vertical lifts. The components of the Hessian g with respect to our standard basis will be denoted as follows:

$$g(\tilde{E}_a,\tilde{E}_b) = g_{ab}, \quad g(X_i,X_j) = g_{ij}, \quad g(X_i,\tilde{E}_a) = g_{ia} = g_{ai} = g(\tilde{E}_a,X_i).$$

We also have $g_{ab}=\tilde{E}_a^V(p_b)$, $g_{ia}=X_i^V(p_a)$. In general, these components are functions on TM, not on M, and the Hessian should be regarded as a tensor field along the tangent bundle projection $\tau: TM\to M$. We will assume throughout that L is regular, which means that g as a whole is nonsingular. Then, Γ is uniquely determined as a second-order differential equation field on TM.

We now turn to the consideration of Routh's procedure. We call the function $\mathcal{R}$ on TM given by

$$\mathcal{R} = L - v^a p_a,$$

the *Routhian*. It generalizes in an obvious way the classical Routhian corresponding to ignorable coordinates. The Routhian is invariant,

$$\tilde{E}_b^C(v^a p_a) = C_{bc}^a v^c p_a - v^a C_{ba}^c p_c = 0,$$

whence the result.

We now consider the Euler–Lagrange equations $\Gamma(X_i^V(L)) - X_i^C(L) = 0$. We wish to write these equations in terms of the restriction of the Routhian to a level set of momentum, say, $p_a = \mu_a$, which we denote by N_μ. To do so, we need to work in terms of vector fields related to X_i^C, X_i^V, and $\tilde{E}_a^C$ which are tangent to N_μ (in general, there is no reason to suppose that these vector fields themselves have this property, of course). To define the new vector fields, we will assume that the Lagrangian has an additional regularity property: we will assume that (g_{ab}) is nonsingular. [Note that if the Hessian is everywhere positive definite, then (g_{ab}) is automatically nonsingular.] Then, there are coefficients A_i^b, B_i^b, and C_a^b, uniquely defined, such that

$$(X_i^C + A_i^b \tilde{E}_b^V)(p_a) = X_i^C(p_a) + A_i^b g_{ab} = 0,$$

$$(X_i^V + B_i^b \tilde{E}_b^V)(p_a) = X_i^V(p_a) + B_i^b g_{ab} = 0,$$

$$(\tilde{E}_a^C + C_a^b \tilde{E}_b^V)(p_c) = \tilde{E}_a^C(p_c) + C_a^b g_{bc} = 0.$$

The vector fields $\bar{X}_i^C$, $\bar{X}_i^V$, and $\bar{E}_a^C$ given by

$$\bar{X}_i^C = X_i^C + A_i^a \tilde{E}_a^V,$$

$$\bar{X}_i^V = X_i^V + B_i^a \tilde{E}_a^V,$$

$$\bar{E}_a^C = \tilde{E}_a^C + C_a^b \tilde{E}_b^V$$

are tangent to each level set N_μ. (The notation is not meant to imply that the barred vector fields are actually complete or vertical lifts.) We will need to know the coefficients explicitly only in the case of B_i^a and C_a^b: in fact,

$$B_i^a = -g^{ab} g_{ib} \quad \text{and} \quad C_a^b = g^{bc} C_{ac}^d p_d.$$

This is all carried out under the assumption that (g_{ab}) is nonsingular. One has to make such an assumption in the classical case in order to be able to solve the equations $\partial L/\partial\dot{\theta}^a = \mu_a$ for the $\dot{\theta}^a$. In the general case, the nonsingularity of (g_{ab}) is the condition for the level set N_μ to be regular, i.e., to define a submanifold of TM of codimension $\dim(\mathfrak{g})$. The vector fields $\tilde{E}_a^V$ are transverse to all regular level sets, and the barred vector fields span the level sets. Thus, on any regular level set, the bracket of any two of the barred vector fields is a linear combination of vector fields of the same form. We want, in particular, to observe that this implies that $[\bar{E}_a^C, \bar{X}_i^V] = 0$. It is not difficult to see, using the known facts about the brackets of the unbarred vector fields, that this bracket is of the form $P^a \tilde{E}_a^V$; this must satisfy $P^a \tilde{E}_a^V(p_b) = P^a g_{ab} = 0$, whence by the regularity assumption $P^a = 0$. In fact, by similar arguments, the brackets of the barred vector fields just reproduce those of their unbarred counterparts, except that $[\bar{X}_i^C, \bar{X}_j^V] = 0$. In particular, $[\bar{E}_a^C, \bar{E}_b^C] = -C_{ab}^c \bar{E}_c^C$. $\bar{E}_a^C$ therefore form an antirepresentation of $\mathfrak{g}$, acting on the level set N_μ (just as $\tilde{E}_a$ do on M).

We return to the expression of the Euler–Lagrange equations in terms of the Routhian. We will need to evaluate the actions of $\bar{X}_i^C$ and $\bar{X}_i^V$ on v^a. Using the formulas in Sec. II we find that

$$\bar{X}_i^C(v^a) = (X_i^C + A_i^b \tilde{E}_b^V)(v^a) = -R_{ij}^a v^j + A_i^a,$$

$$\bar{X}_i^V(v^a) = (X_i^V + B_i^b \tilde{E}_b^V)(v^a) = B_i^a.$$

We now set things up so that we can restrict to the submanifold N_μ easily. We have

$$X_i^C(L) = \bar{X}_i^C(L) - A_i^a \tilde{E}_a^V(L) = \bar{X}_i^C(L - v^a p_a) + (-R_{ij}^a v^j + A_i^a)p_a + v^a \bar{X}_i^C(p_a) - A_i^a p_a = \bar{X}_i^C(\mathcal{R}) - p_a R_{ij}^a v^j,$$

$$X_i^V(L) = \bar{X}_i^V(L) - B_i^a \tilde{E}_a^V(L) = \bar{X}_i^V(L - v^a p_a) + B_i^a p_a + v^a \bar{X}_i^V(p_a) - B_i^a p_a = \bar{X}_i^V(\mathcal{R}).$$

However, $\Gamma(X_i^V(L)) - X_i^C(L) = 0$, and Γ is tangent to the submanifold N_μ; thus, if we denote by $\mathcal{R}^\mu$ the restriction of the Routhian to the submanifold (where it becomes $L - v^a \mu_a$), we have

$$\Gamma(\bar{X}_i^V(\mathcal{R}^\mu)) - \bar{X}_i^C(\mathcal{R}^\mu) = -\mu_a R_{ij}^a v^j.$$

On the other hand, if Γ is a second-order differential equation field such that $\Gamma(\tilde{E}_a^V(L)) = 0$ and the above equation holds for all μ_a, then Γ satisfies the Euler–Lagrange equations for the invariant Lagrangian L.

We will refer to these equations as the generalized Routh equations. Neither $\mathcal{R}^\mu$ nor Γ is $\bar{E}_a^C$-invariant. They will, however, be invariant under those vector fields $\tilde{\xi}^C$, $\xi \in \mathfrak{g}$, which happen to be tangent to the level set N_μ. These are the vector fields for which $\xi^a \tilde{E}_a^C = \xi^a \bar{E}_a^C$, or $\xi^a C_{ab}^c \mu_c = 0$. We will return to this issue in later sections.

Note that since Γ satisfies $\Gamma(p_a) = 0$, it may be expressed in the form

$$\Gamma = v^i \bar{X}_i^C + \Gamma^i \bar{X}_i^V + v^a \bar{E}_a^C.$$

If the matrix-valued function $\bar{X}_i^V \bar{X}_j^V(\mathcal{R})$ is nonsingular, the reduced Euler–Lagrange equations above will determine the coefficients Γ^i. We show now that this is always the case, under the assumptions made earlier.

Recall that $\bar{X}_i^V = X_i^V + B_i^a \tilde{E}_a^V$ is determined by the condition that $\bar{X}_i^V(p_a) = 0$ and that therefore $B_i^a = -g^{ab} g_{ib}$. We may regard $X_i + B_i^a \tilde{E}_a$ as a vector field along the tangent bundle projection, and $\bar{X}_i^V$ really is the vertical lift of this vector field; we will accordingly denote it by $\bar{X}_i$. Then,

$$g(\bar{X}_i, \tilde{E}_a) = g(X_i, \tilde{E}_a) + B_i^b g(\tilde{E}_b, \tilde{E}_a) = g_{ia} + B_i^b g_{ab} = 0.$$

Thus, $\bar{X}_i$ span the orthogonal complement to the space spanned by $\tilde{E}_a$ with respect to the Hessian of L. That is to say, the tangent space to a regular level set of momentum at any point $u \in TM$ intersects the tangent space to the fiber of $TM \to M$ at u in the subspace orthogonal with respect to g_u to the span of the $\tilde{E}_a^V$. Moreover,

$$g(\bar{X}_i, \bar{X}_j) = g_{ij} + B_i^a g_{aj} + B_j^a g_{ia} + B_i^a B_j^b g_{ab} = g_{ij} - 2g^{ab} g_{ia} g_{jb} + g^{ac} g^{bd} g_{ic} g_{jd} g_{ab} = g_{ij} - g^{ab} g_{ia} g_{jb}.$$

So, this is the expression for the restriction of the Hessian of L to the subspace orthogonal to that spanned by the $\tilde{E}_a$.

Now, recall that $\bar{X}_i^V(\mathcal{R}) = X_i^V(L)$. Thus,

$$\bar{X}_i^V \bar{X}_j^V(\mathcal{R}) = (X_i^V - g^{ab} g_{ib} \tilde{E}_a^V) X_j^V(L) = g_{ij} - g^{ab} g_{ib} g_{aj} = g(\bar{X}_i, \bar{X}_j).$$

That is, the "Hessian" of $\mathcal{R}$ [i.e., $\bar{X}_i^V \bar{X}_j^V(\mathcal{R})$] is just the restriction of the Hessian of L to the subspace orthogonal to that spanned by $\tilde{E}_a$. It follows that the bilinear form with components $\bar{g}_{ij} = \bar{X}_i^V \bar{X}_j^V(\mathcal{R})$ must be nonsingular. Suppose that there is some vector w^j such that $\bar{g}_{ij} w^j = 0$; then, $g(\bar{X}_i, w^j \bar{X}_j) = 0$ by assumption and $g(\tilde{E}_a, w^j \bar{X}_j) = 0$ by orthogonality—but then $w^j \bar{X}_j = 0$ since g is assumed to be nonsingular.

The sense in which the generalized Routh equations are "reduced" Euler–Lagrange equations is that (in principle at least) we can reduce the number of variables by using the equations $p_a = \mu_a$ to eliminate the quasivelocities v^a. However, these variables appear explicitly in the expression for Γ, so it may be considered desirable to rearrange the generalized Routh equations so that they no longer appear. This can be done by changing the basis of vector fields on the level set of momentum as follows. The change is suggested by the fact that, notation notwithstanding, $S(\bar{X}_i^C) \neq \bar{X}_i^V$ (where S is the vertical endomorphism). Let us, however, set

$$\hat{X}_i^C = \bar{X}_i^C + B_i^a \bar{E}_a^C.$$

Then, since S vanishes on vertical lifts,

$$S(\hat{X}_i^C) = S(X_i^C + B_i^a \tilde{E}_a^C) = X_i^V + B_i^a \tilde{E}_a^V = \bar{X}_i^V.$$

We write

$$\Gamma_0 = v^i \hat{X}_i^C + \Gamma^i \bar{X}_i^V,$$

so that

$$\Gamma = \Gamma_0 + (v^i B_i^a + v^a)\bar{E}_a^C.$$

We will examine the contribution of the term involving $\bar{E}_a^C$ in Γ to the generalized Routh equations. First, we determine $\bar{E}_a^C(\mathcal{R})$. Since $\bar{E}_a^C(p_b)=0$,

$$\bar{E}_a^C(\mathcal{R}) = \bar{E}_a^C(L - v^b p_b) = C_a^b p_b - \bar{E}_a^C(v^b) p_b = C_a^b p_b - C_{ac}^b p_b v^c - C_a^c \delta_c^b p_b = -C_{ac}^b p_b v^c.$$

It follows that

$$\bar{E}_a^C(\bar{X}_i^V(\mathcal{R})) = \bar{X}_i^V(\bar{E}_a^C(\mathcal{R})) = -\bar{X}_i^V(C_{ac}^b p_b v^c) = -C_{ac}^b p_b B_i^c.$$

So, setting $\Gamma = \Gamma_0 + (v^i B_i^a + v^a)\bar{E}_a^C$, we have

$$\begin{aligned}\Gamma(\bar{X}_i^V(\mathcal{R}^\mu)) - \bar{X}_i^C(\mathcal{R}^\mu) &= \Gamma_0(\bar{X}_i^V(\mathcal{R}^\mu)) + (v^j B_j^a + v^a)\bar{E}_a^C(\bar{X}_i^V(\mathcal{R}^\mu)) - \hat{X}_i^C(\mathcal{R}^\mu) + B_i^a \bar{E}_a^C(\mathcal{R}^\mu)\\ &= \Gamma_0(\bar{X}_i^V(\mathcal{R}^\mu)) - \hat{X}_i^C(\mathcal{R}^\mu) - (v^j B_j^a + v^a) C_{ac}^b \mu_b B_i^c - B_i^a C_{ac}^b \mu_b v^c = \Gamma_0(\bar{X}_i^V(\mathcal{R}^\mu))\\ &\quad - \hat{X}_i^C(\mathcal{R}^\mu) + v^j B_j^a B_i^c C_{ac}^b \mu_b,\end{aligned}$$

and the generalized Routh equations become

$$\Gamma_0(\bar{X}_i^V(\mathcal{R}^\mu)) - \hat{X}_i^C(\mathcal{R}^\mu) = -\mu_a(R_{ij}^a + B_i^b B_j^c C_{bc}^a) v^j.$$

We may say that among the vector fields tangent to a level set of momentum, it is $\bar{X}_i^C + B_i^a \bar{E}_a^C$, not $\bar{X}_i^C$, that really plays the role of the complete lift of $\bar{X}_i$. Be aware, however, that unless the symmetry group is Abelian, Γ_0 cannot be identified with a vector field on $T(M/G)$. We will return to this matter at the end of Sec. V.

To end this section, we give a coordinate expression for the generalized Routh equations in their original form. For this purpose, we take coordinates (x^i) on M/G and coordinates (x^i, θ^a) on M such that the θ^a are fiber coordinates; then, (x^i, θ^a, v^i) are coordinates on N_μ, which is to say that N_μ can be locally identified with $M \times_{M/G} T(M/G)$. We may write

$$X_i = \frac{\partial}{\partial x^i} - \Lambda_i^a \frac{\partial}{\partial \theta^a}, \quad \tilde{E}_a = K_a^b \frac{\partial}{\partial \theta^b},$$

for suitable functions Λ_i^a and K_a^b on M. (We should note that K_a^b are components of a nonsingular matrix at each point; moreover, the invariance property of the X_i can be expressed in terms of the coefficients Λ_i^a and K_a^b but we will not actually need either of these facts here.) From the formulas for the action of complete and vertical lifts on quasivelocities given at the end of Sec. II, we see that

$$\bar{X}_i^C(v^i) = \bar{E}_a^C(v^i) = 0, \quad \bar{X}_i^V(v^j) = \delta_i^j.$$

Thus, in terms of x^i, θ^a, and v^i we can write

$$\bar{X}_i^C = \frac{\partial}{\partial x^i} - \Lambda_i^a \frac{\partial}{\partial \theta^a}, \quad \bar{X}_i^V = \frac{\partial}{\partial v^i}, \quad \bar{E}_a^C = K_a^b \frac{\partial}{\partial \theta^b}.$$

It is necessary to be a little careful: the coordinate vector field expressions are ambiguous since they can refer either to coordinates on TM or on N_μ. We emphasize that it is the latter interpretation that is intended here. In view of the possibilities of confusion, it will be useful to have an explicit notation for the injection $N_\mu \to TM$: we denote it by ι. The nonsingularity of (g_{ab}) ensures that, at least locally, we can rewrite the relation $p_a = \mu_a$ for the injection $\iota: N_\mu \to TM$ in the form $v^a = \iota^a(x^i, \theta^a, v^i)$, for certain functions ι^a of the specified variables.

The restriction of the Euler–Lagrange field Γ to N_μ is

$$\Gamma = \iota^a \bar{E}_a^C + v^i \bar{X}_i^C + (\Gamma^i \circ \iota)\bar{X}_i^V = \iota^b K_b^a \frac{\partial}{\partial \theta^a} + v^i\left(\frac{\partial}{\partial x^i} - \Lambda_i^a \frac{\partial}{\partial \theta^a}\right) + (\Gamma^i \circ \iota)\frac{\partial}{\partial v^i} = (\iota^b K_b^a - v^i \Lambda_i^a)\frac{\partial}{\partial \theta^a} + v^i \frac{\partial}{\partial x^i} + (\Gamma^i \circ \iota)\frac{\partial}{\partial v^i}.$$

The equations for its integral curves are

$$\dot{x}^i = v^i,$$

$$\dot{v}^i = \Gamma^i(x, \theta, v),$$

$$\dot{\theta}^a = \iota^b(x, \theta, v) K_b^a(x, \theta) - v^i \Lambda_i^a(x, \theta).$$

These can be considered as a coupled set of first- and second-order equations,

$$\ddot{x}^i = \Gamma^i(x, \theta, \dot{x}),$$

$$\dot{\theta}^a = \iota^b(x, \theta, \dot{x}) K_b^a(x, \theta) - \dot{x}^i \Lambda_i^a(x, \theta).$$

With regard to the second of these equations, we point out that the expression for the velocity variables $\dot{\theta}^a$ in terms of the quasivelocities v^i and v^a is just $\dot{\theta}^a = v^b K_b^a - v^i \Lambda_i^a$. What turns these identities into genuine differential equations is, in particular, the substitution for the v^a in terms of the other variables via the functions ι^a—or, in other words, restriction to N_μ.

The functions Γ^i may be determined from the generalized Routh equations. These may be expressed as

$$\frac{d}{dt}\left(\frac{\partial \mathcal{R}^\mu}{\partial v^i}\right) - \frac{\partial \mathcal{R}^\mu}{\partial x^i} = -\mu_a R^a_{ij} - \Lambda^a_i \frac{\partial \mathcal{R}^\mu}{\partial \theta^a}.$$

In the light of the earlier remarks about the interpretation of coordinate vector fields, we point out that substitution for v^a in terms of the other variables in this equation must be carried out before the partial derivatives are calculated.

IV. THE RECONSTRUCTION METHOD

We have seen in Sec. III that Routh's technique consists in restricting the Euler–Lagrange equations to a level set of momentum N_μ. This procedure takes partial, but not necessarily complete, account of the action of the symmetry group G. To make further progress, we must examine the residual action of G on N_μ.

As we mentioned before, the momentum map is equivariant between the induced action of G on TM and the coadjoint action of G on $\mathfrak{g}^*$. The submanifold N_μ is therefore invariant under the isotropy group $G_\mu = \{g \in G \,|\, \mathrm{ad}^*_g\, \mu = \mu\}$ of μ. The algebra $\mathfrak{g}_\mu$ of G_μ consists of those $\xi \in \mathfrak{g}$ such that $\xi^b C^c_{ab}\mu_c = 0$; this is the necessary and sufficient condition for $\tilde{\xi}^C$ to be tangent to N_μ. Note that any geometric object we know to be G-invariant is automatically G_μ-invariant.

The manifold N_μ is a principal fiber bundle with group G_μ; we will denote its base by N_μ / G_μ. The restriction of the Euler–Lagrange field Γ to N_μ is G_μ-invariant and, as a consequence, it projects onto a vector field $\check{\Gamma}$ on N_μ / G_μ.

The task now is to examine the relationship between $\check{\Gamma}$ and Γ. There are two aspects: the formulation of the differential equations represented by $\check{\Gamma}$ and the reconstruction of integral curves of Γ from integral curves of $\check{\Gamma}$ (supposing that we have solved those equations).

Our methods of attack on these problems will be based on those we developed in our papers[3,11] and are similar to (but different from) the ones that were adopted in, e.g., Ref. 7. These in turn were based on the following well-known method for reconstructing integral curves of an invariant vector field from reduced data. Let $\pi: N \to B$ be a principal fiber bundle with group G. Any invariant vector field Γ on N defines a π-related reduced vector field $\check{\Gamma}$ on B: due to the invariance of Γ, the relation $T\pi(\Gamma(n)) = \check{\Gamma}(\pi(n))$ is independent of the choice of $n \in N$ within the equivalence class of $\pi(n) \in B$. Given a principal connection Ω, an integral curve $v(t)$ of Γ can be reconstructed from an integral curve $\check{v}(t)$ of $\check{\Gamma}$ as follows. Let $\check{v}^H(t)$ be a horizontal lift of $\check{v}(t)$ with respect to Ω [that is, a curve in N over $\check{v}$ such that $\Omega(\check{v}^H) = 0$] and let $g(t)$ be the solution in G of the equation

$$\vartheta(\tilde{\dot{g}}(t)) = \Omega(\Gamma(v^H(t))),$$

where ϑ is the Maurer–Cartan form of G. [We use here the fact that given any curve $\xi(t)$ in $\mathfrak{g}$, the Lie algebra of G, there is a unique curve $g(t)$ in G which satisfies $\vartheta(\dot{g}(t)) = \xi(t)$ and $g(0) = e$; $g(t)$ is sometimes called the development of $\xi(t)$ into G, see, for example Ref. 14.] Then, $v(t) = \psi^N_{g(t)} v^H(t)$ is an integral curve of Γ.

In the following sections, we define two principal connections on N_μ, we determine Γ, and we identify for both connections the vertical part of Γ, necessary for the reconstruction method above.

V. TWO PRINCIPAL CONNECTIONS ON A LEVEL SET OF MOMENTUM

A principal connection Ω on $N_\mu \to N_\mu / G_\mu$ is by definition a left splitting of the short exact sequence

$$0 \to N_\mu \times \mathfrak{g}_\mu \to TN_\mu \to N_\mu \times_{N_\mu/G_\mu} T(N_\mu/G_\mu) \to 0.$$

All spaces in the above sequence should be interpreted as bundles over N_μ. We think of Ω as a type (1,1) tensor field on N_μ which is pointwise a projection operator with image the tangent space to the fiber, and which is invariant under G_μ.

The first connection we define uses the Hessian of L to determine its horizontal distribution and is therefore analogous to the mechanical connection of a simple system; we denote it by Ω^m. Recall that we interpret the Hessian g of L as a tensor field along τ. In particular, its components with respect to the standard basis $\{X_i, \widetilde{E}_a\}$ are functions on TM. We will say that a vector field W on N_μ is horizontal for Ω^m if

$$g(\widetilde{\xi}, \tau_* W) = 0, \quad \forall\, \xi \in \mathfrak{g}_\mu,$$

where $\tau_* W$ is the projection of a vector field W on TM to a vector field along $\tau: TM \to M$. The definition makes sense only if we assume that the restriction of g to $N_\mu \times \mathfrak{g}_\mu$ is nonsingular, as we do from now on.

In Ref. 11, we have shown that if the Lagrangian is invariant, then so is g, in the sense that

$$\mathcal{L}_{\widetilde{\xi}}\, g = 0, \quad \forall\, \xi \in \mathfrak{g}.$$

Here, for a vector field Z on M, $\mathcal{L}_Z$ stands for an operator acting on tensor fields along τ that has all the properties of a Lie derivative operator and, in particular, when applied to a function f on TM and a vector field X along τ gives

$$\mathcal{L}_Z f = Z^C(f), \quad \mathcal{L}_Z X = \left(Z^\beta \frac{\partial X^\alpha}{\partial x^\beta} + \frac{\partial Z^\beta}{\partial x^\gamma} u^\gamma \frac{\partial X^\alpha}{\partial u^\beta} - X^\beta \frac{\partial Z^\alpha}{\partial x^\beta} \right) \frac{\partial}{\partial x^\alpha},$$

where

$$Z = Z^\alpha(x) \frac{\partial}{\partial x^\alpha}, \quad X = X^\alpha(x,u) \frac{\partial}{\partial x^\alpha}.$$

Note that if X is a basic vector field along τ (i.e., a vector field on M), then $\mathcal{L}_Z X = [Z, X]$. Furthermore, for any vector field W on TM, we have

$$\mathcal{L}_Z(\tau_* W) = \tau_*[Z^C, W].$$

To show that the connection is principal, we need only to show that if W is horizontal, so also is $[\widetilde{\xi}^C, W]$ for all $\xi \in \mathfrak{g}_\mu$. However, for all $\xi, \eta \in \mathfrak{g}_\mu$,

$$g(\tau_*[\widetilde{\xi}^C, W], \widetilde{\eta}) = g(\mathcal{L}_{\widetilde{\xi}}(\tau_* W), \widetilde{\eta}) = -g(\tau_* W, \mathcal{L}_{\widetilde{\xi}} \widetilde{\eta}) = -g(\tau_* W, [\widetilde{\xi}, \widetilde{\eta}]) = g(\tau_* W, \widetilde{[\xi, \eta]}) = 0,$$

using the properties of the generalized Lie derivative and the invariance of g.

As was mentioned before, $\{\bar{X}_i^C, \bar{E}_a^C, \bar{X}_i^V\}$ is a basis of vector fields on N_μ. Suppose now that the basis $\{E_a\} = \{E_A, E_\alpha\}$ of $\mathfrak{g}$ is chosen so that $\{E_A\}$ is a basis of $\mathfrak{g}_\mu$. Then, $C_{Ab}^c \mu_c = 0$, and on N_μ, we get for the corresponding fundamental vector fields

$$\bar{E}_A^C = \widetilde{E}_A^C + g^{bc} C_{Ac}^d \mu_d \widetilde{E}_b^V = \widetilde{E}_A^C.$$

All $\widetilde{E}_A^C$ are therefore tangent to N_μ, as required. These vector fields span exactly the vertical space of $N_\mu \to N_\mu / G_\mu$ which we have identified with $N_\mu \times \mathfrak{g}_\mu$. Vector fields of this form are infinitesimal generators of the G_μ-action on N_μ.

If (G^{AB}) is the inverse of the matrix (g_{AB}) [and not the (A,B)-component of (g^{ab})], then the vector fields,

$$\bar{E}_\alpha^H = \bar{E}_\alpha^C - G^{AB} g_{A\alpha} \widetilde{E}_B^C = \bar{E}_\alpha^C - \Upsilon_\alpha^B \widetilde{E}_B^C,$$

$$\bar{X}_i^H = \bar{X}_i^C - G^{AB} g_{Ai} \tilde{E}_B^C = \bar{X}_i^C - \Upsilon_i^B \tilde{E}_B^C,$$

together with $\bar{X}_i^V$, are horizontal. (As was the case with the notations $\bar{E}_a^C$, etc., the notation for the horizontal fields is not meant to imply that $\bar{E}_\alpha^H$, etc., are actually horizontal lifts.) The action of Ω^m is simply

$$\Omega^m(\tilde{E}_A^C) = \tilde{E}_A^C, \quad \Omega^m(\bar{E}_\alpha^H) = 0, \quad \Omega^m(\bar{X}_i^H) = 0, \quad \Omega^m(\bar{X}_i^V) = 0,$$

and since the arguments form a basis of vector fields on N_μ, these equations specify Ω^m explicitly. We will call Ω^m the mechanical connection on N_μ.

The vector fields $\hat{X}_i^C = \bar{X}_i^C + B_i^a \bar{E}_a^C$ introduced earlier are also horizontal; they can be expressed as $\hat{X}_i^C = \bar{X}_i^H - g^{\alpha b} g_{bi} \bar{E}_\alpha^H$.

The vector fields $\bar{X}_i^C$ are not horizontal with respect to Ω^m. However, it is possible to identify a second principal connection Ω^{N_μ} on N_μ for which these vector fields are horizontal. We will identify Ω^{N_μ} in two steps.

It seems natural to split the basis $\{\bar{X}_i^C, \bar{E}_a^C, \bar{X}_i^V\}$ into a "vertical" part $\{\bar{E}_a^C\}$ and a "horizontal" part $\{\bar{X}_i^C, \bar{X}_i^V\}$. To see that it does indeed make sense to do so, it is sufficient to observe that the distributions spanned by $\{\bar{E}_a^C\}$ and $\{\bar{X}_i^C, \bar{X}_i^V\}$, respectively, are unchanged when the bases $\{E_a\}$ of $\mathfrak{g}$ and $\{X_i\}$ of ω-horizontal vector fields on M are replaced by different ones. Under a change of basis for $\mathfrak{g}$, $\{\bar{E}_a^C\}$ are simply replaced by constant linear combinations of themselves, so their span is clearly unchanged. On the other hand, if we set $Y_i = A_i^j X_j$, then $Y_i^V = A_i^j X_j^V$ and $Y_i^C = A_i^j X_j^C + \dot{A}_i^j X_j^V$ (where $\dot{A}_i^j$ is the total derivative of A_i^j, not that it matters), so the distributions spanned by $\{\bar{X}_i^C, \bar{X}_i^V\}$ and $\{\bar{Y}_i^C, \bar{Y}_i^V\}$ are the same.

So, we can indeed characterize a connection in this way, but it is not a connection on $N_\mu \to N_\mu / G_\mu$. In fact, this construction defines a connection on the bundle with projection $N_\mu \to T(M/G)$ [the restriction of $T\pi^M : TM \to T(M/G)$ to N_μ], i.e., a splitting of the short exact sequence

$$0 \to N_\mu \times \mathfrak{g} \to TN_\mu \to N_\mu \times_{T(M/G)} T(T(M/G)) \to 0.$$

[Recall that the vector fields $\bar{E}_a^C$, which span the vertical space of the projection $N_\mu \to T(M/G)$, form an antirepresentation of $\mathfrak{g}$ acting on the level set N_μ.] The construction just described is a version of the so-called vertical lift of a connection on a principal bundle (here ω) to its tangent bundle (this is described more fully in Ref. 3); accordingly, we denote the corresponding type (1,1) tensor field by Ω^V, and we have

$$\Omega^V(\bar{E}_a^C) = \bar{E}_a^C, \quad \Omega^V(\bar{X}_i^C) = 0, \quad \Omega^V(\bar{X}_i^V) = 0.$$

Evidently, $(\Omega^V)^2 = \Omega^V$. We show now that $\mathcal{L}_{\tilde{E}_A^C} \Omega^V = 0$ for all A. Firstly, note that

$$[\tilde{E}_A^C, \bar{E}_a^C] = [\bar{E}_A^C, \bar{E}_a^C] = -C_{Aa}^b \bar{E}_b^C,$$

so that

$$(\mathcal{L}_{\tilde{E}_A^C} \Omega^V)(\bar{E}_a^C) = [\tilde{E}_A^C, \Omega^V(\bar{E}_a^C)] - \Omega^V[\tilde{E}_A^C, \bar{E}_a^C] = [\tilde{E}_A^C, \bar{E}_a^C] + C_{Aa}^b \bar{E}_b^C = 0.$$

Moreover, since $[\tilde{E}_A^C, \bar{X}_i^C] = 0$,

$$(\mathcal{L}_{\tilde{E}_A^C} \Omega^V)(\bar{X}_i^C) = [\tilde{E}_A^C, \Omega^V(\bar{X}_i^C)] - \Omega^V[\tilde{E}_A^C, \bar{X}_i^C] = 0,$$

and similarly for $\bar{X}_i^V$.

The relation between Ω^V and ω may be described more easily if we momentarily break our convention by specifying connections by their forms rather than by the tensors corresponding to their splittings: it is easily checked that

$$\Omega^V(Z_v) = \omega(T(\tau \circ \iota)Z_v), \quad Z_v \in TN_\mu.$$

This equation has to be read as one between elements of $\mathfrak{g}$, obtained by identifying the vertical sub-bundle of TN_μ with $N_\mu \times \mathfrak{g}$ and the vertical subbundle of TN with $M \times \mathfrak{g}$, or if you will by projection onto $\mathfrak{g}$.

The vertical space $N_\mu \times \mathfrak{g}_\mu$ of the connection Ω^{N_μ} we are looking for is only a sub-bundle of the vertical space $N_\mu \times \mathfrak{g}$ of the connection Ω^V. So, in a second step, we need to identify a connection for the following sequence of trivial vector bundles:

$$0 \to N_\mu \times \mathfrak{g}_\mu \to N_\mu \times \mathfrak{g} \to N_\mu \times \mathfrak{g}/\mathfrak{g}_\mu \to 0.$$

For this connection, we can simply take the restriction of the mechanical connection Ω^m defined earlier to the submanifold $N_\mu \times \mathfrak{g}$. The connection Ω^{N_μ} is then simply $\Omega^m \circ \Omega^V$ (see the diagram below).

$$\begin{tikzcd}
N_\mu \times \mathfrak{g}_\mu \arrow[r] \arrow[d] & N_\mu \times \mathfrak{g} \arrow[r] \arrow[d] & N_\mu \times \mathfrak{g}/\mathfrak{g}_\mu \arrow[d] \\
N_\mu \times \mathfrak{g}_\mu \arrow[r] \arrow[d] & TN_\mu \arrow[r] \arrow[d] & N_\mu \times_{N_\mu/G_\mu} T(N_\mu/G_\mu) \arrow[d] \\
0 \arrow[r] & N_\mu \times_{T(M/G)} T(T(M/G)) \arrow[r] & N_\mu \times_{T(M/G)} T(T(M/G))
\end{tikzcd}$$

By construction, $\Omega^V \circ \Omega^m = \Omega^m$, so $\Omega^{N_\mu} = \Omega^m \circ \Omega^V$ satisfies $(\Omega^{N_\mu})^2 = \Omega^{N_\mu}$ as it should. We have

$$\Omega^{N_\mu}(\tilde{E}_A^C) = \tilde{E}_A^C, \quad \Omega^{N_\mu}(\bar{E}_\alpha^H) = 0, \quad \Omega^{N_\mu}(\bar{X}_i^C) = 0, \quad \Omega^{N_\mu}(\bar{X}_i^V) = 0.$$

The tensor field Ω^{N_μ} is G_μ-invariant since both of the tensors of which it is composed are G_μ-invariant; Ω^{N_μ} therefore defines a principal G_μ-connection.

Note that to define the mechanical connection, we do not need a principal connection ω on $M \to M/G$ (though we may use one in calculations). If such a connection is available, then we can use either Ω^m or Ω^{N_μ} for the reconstruction method.

The connection Ω^{N_μ} is clearly different from Ω^m in general. We can also decompose Ω^m into two connections, in accordance with the short exact sequences in the diagram. The splitting Ω_0^V of the middle vertical line, similar to the connection Ω^V of Ω^{N_μ}, can be defined by saying that a vector field W is horizontal if $g(\tilde{\xi}, \tau_* W) = 0$ for all $\xi \in \mathfrak{g}$ (not just for $\xi \in \mathfrak{g}_\mu$). For this connection, we have

$$\Omega_0^V(\bar{E}_a^C) = \bar{E}_a^C, \quad \Omega_0^V(\hat{X}_i^C) = 0, \quad \Omega^V(\bar{X}_i^V) = 0,$$

where the vector fields $\hat{X}_i^C$ are exactly those that we have encountered in Sec. II.

To end this section, we consider the decomposition of the restriction of the Euler–Lagrange field Γ to N_μ into its vertical and horizontal parts with respect to the two connections.

Let us introduce coordinates (x^i, θ^a) on M such that the orbits of G (or, in other words, the fibers of $M \to M/G$) are given by $x^i = \text{const}$; x^i may therefore be regarded as coordinates on M/G. As before, we will use as fiber coordinates the quasivelocities (v^i, v^a) with respect to the standard basis $\{X_i, \widetilde{E}_a\}$. The nonsingularity of (g_{ab}) ensures that, at least locally, we can rewrite the relation $p_a=\mu_a$ for the injection $\iota: N_\mu \to TM$ in the form $v^a=\iota^a(x^i,\theta^a,v^i)$, for certain functions ι^a of the specified variables. The restriction of the Euler–Lagrange field to N_μ is

$$\Gamma = \iota^a \bar{E}_a^C + v^i \bar{X}_i^C + (\Gamma^i \circ \iota)\bar{X}_i^V = (\iota^A + \Upsilon_\alpha^A \iota^\alpha)\widetilde{E}_A^C + \iota^\alpha \bar{E}_\alpha^H + v^i \bar{X}_i^C + (\Gamma^i \circ \iota)\bar{X}_i^V = (\iota^A + \Upsilon_\alpha^A \iota^\alpha + \Upsilon_i^A v^i)\widetilde{E}_A^C + \iota^\alpha \bar{E}_\alpha^H + v^i \bar{X}_i^H + (\Gamma^i \circ \iota)\bar{X}_i^V.$$

The vertical part of Γ with respect to the mechanical connection Ω^m is $(\iota^A+\Upsilon_\alpha^A\iota^\alpha+\Upsilon_i^A v^i)\widetilde{E}_A^C$, and with respect to the vertical lift connection Ω^{N_μ}, it is $(\iota^A+\Upsilon_\alpha^A\iota^\alpha)\widetilde{E}_A^C$.

Note that neither of the current decompositions of Γ coincides with the one we had toward the end of Sec. III, which we should now write $\Gamma=(\iota^a+B_i^a v^i)\bar{E}_a^C+\Gamma_0$. The reason is that this last decomposition is only partial, in the sense that the vector field Γ_0 is the horizontal part of Γ with respect to the connection Ω_0^V; it is the horizontal lift of a section of the pullback bundle $N_\mu \times_{T(M/G)} T(T(M/G))$, not a vector field on $T(M/G)$, and this section is only a part of the data required for the reconstruction method.

VI. THE REDUCED VECTOR FIELD

A principal connection is all we need to reconstruct integral curves of an invariant vector field from those of its reduced vector field. We next examine the latter.

A. The Abelian case

Before embarking on the more general case, it is instructive to see what happens if the symmetry group G happens to be Abelian, i.e., when $C_{ab}^c=0$. Then, as we pointed out earlier for the case of a simple mechanical system with Abelian symmetry group, $\mathfrak{g}_\mu=\mathfrak{g}$ and any level set $p_a=\mu_a$ is invariant under the whole group G. In fact, under the assumption that $p_a=\mu_a$ can be solved locally in the form $v^a=\iota^a$, N_μ/G can be interpreted as $T(M/G)$, with coordinates (x^i,v^i), where x^i are coordinates on M/G and v^i the corresponding fiber coordinates (no longer *quasivelocities*). In this case, there are no E_α-vectors and $\bar{E}_a^C=\widetilde{E}_a^C$ for all a.

The restriction of the Euler–Lagrange field to N_μ, given here by

$$\Gamma = \iota^a \widetilde{E}_a^C + v^i \bar{X}_i^C + (\Gamma^i \circ \iota)\bar{X}_i^V = (\iota^a - g^{ab} g_{bi} v^i)\widetilde{E}_a^C + v^i \bar{X}_i^H + (\Gamma^i \circ \iota)\bar{X}_i^V,$$

is now also G-invariant. As a consequence, the coefficients $\Gamma^i \circ \iota$ do not depend on the group coordinates θ^a but only on the coordinates (x^i,v^i) of $T(M/G)$. In fact, the vector fields $v^i\bar{X}_i^C+(\Gamma^i\circ\iota)\bar{X}_i^V$ (the Ω^{N_μ}-horizontal part of Γ) and $v^i\bar{X}_i^H+(\Gamma^i\circ\iota)\bar{X}_i^V$ (the Ω^m-horizontal part of Γ) both reduce to the same vector field on N_μ/G, which in this case is exactly a second-order differential equation field on $T(M/G)$,

$$\Gamma = v^i \frac{\partial}{\partial x^i} + \Gamma^i(x,v)\frac{\partial}{\partial v^i}.$$

The integral curves of this reduced vector field are the solutions of the equations $\ddot{x}^i=\Gamma^i(x,\dot{x})$ (with $v^i=\dot{x}^i$) and, from the introduction, we know that these are equivalent to the equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{R}}{\partial v^i}\right) - \frac{\partial \mathcal{R}}{\partial x^i} = -B_{ij}^a \pi_a \dot{x}^j.$$

B. The non-Abelian case

In the general case of a non-Abelian symmetry group, we should not expect that the equations for x^i will be completely decoupled from all coordinates θ^a. Indeed, in that case, the vector field

Γ reduces to a vector field on N_μ/G_μ. This manifold can locally be identified with $M/G_\mu \times T(M/G)$, so the equations for the integral curves of the reduced vector field will depend also on the coordinates of M/G_μ.

To give a local expression of the reduced vector field $\check{\Gamma}$, we need to introduce a basis for $\mathcal{X}(N_\mu/G_\mu)$. The bracket relations $[\tilde{E}_A^C,\bar{X}_i^C]=0$ and $[\tilde{E}_A^C,\bar{X}_i^V]=0$ show that $\bar{X}_i^C$ and $\bar{X}_i^V$ are G_μ-invariant vector fields on N_μ; they project therefore onto vector fields X_i^C and X_i^V on N_μ/G_μ. The invariance of the Hessian g amounts for its coefficients to $\tilde{E}_a^C(g_{ij})=0$, $\tilde{E}_a^C(g_{bc})+C_{ab}^d g_{cd}+C_{ac}^d g_{bd}=0$, and $\tilde{E}_a^C(g_{ib})+C_{ab}^c g_{ic}=0$. From this,

$$\tilde{E}_A^C(\Upsilon_i^B)=C_{AC}^B\Upsilon_i^C,$$

$$\tilde{E}_A^C(\Upsilon_\alpha^B)=C_{AC}^B\Upsilon_\alpha^C-C_{A\alpha}^\beta\Upsilon_\beta^B-C_{A\alpha}^B$$

(where we have taken into account the fact that in the current basis $C_{AB}^\gamma=0$). It is now easy to see that the vector fields $\bar{X}_i^H=\bar{X}_i^C-\Upsilon_i^A\tilde{E}_A^C$ are also G_μ-invariant. In fact, since they differ from $\bar{X}_i^C$ only in a part that is vertical with respect to the bundle projection $N_\mu\to N_\mu/G_\mu$, they project onto the same vector fields $\check{X}_i^C$ on N_μ/G_μ.

The vector fields $\bar{E}_\alpha^H$ are not invariant: in fact, $[\tilde{E}_A^C,\bar{E}_\alpha^H]=-C_{A\alpha}^\beta\bar{E}_\beta^H$. To obtain a complete basis for $\mathcal{X}(N_\mu/G_\mu)$, we need to replace the vector fields $\{\bar{E}_\alpha^H\}$ by G_μ-invariant vector fields. To do so, we will consider the G-invariant vector fields $\hat{E}_a=\mathcal{A}_a^b\tilde{E}_b$ on M that we introduced in Sec. II. Let $(\mathcal{A}_\beta^\alpha)$ be the coefficients we find in the relation $\hat{E}_\alpha=\mathcal{A}_\alpha^\beta\tilde{E}_\beta+\mathcal{A}_\alpha^B\tilde{E}_B$. The vector fields

$$\hat{E}_\alpha^H=\mathcal{A}_\alpha^\beta\bar{E}_\beta^H$$

are tangent to the level set N_μ and horizontal. Given that $C_{AB}^\beta=0$, it easily follows from the relation $\tilde{E}_A(\mathcal{A}_\alpha^\beta)=C_{A\gamma}^\beta\mathcal{A}_\alpha^\gamma$ that these vector fields are G_μ-invariant,

$$[\tilde{E}_A^C,\hat{E}_\alpha^H]=\tilde{E}_A(\mathcal{A}_\alpha^\beta)\bar{E}_\beta^H-\mathcal{A}_\alpha^\beta C_{A\beta}^\gamma\bar{E}_\gamma^H=0.$$

They project therefore onto vector fields $\check{E}_\alpha^H$ on N_μ/G_μ. To conclude, the set $\{\check{X}_i^C,\check{X}_i^V,\check{E}_\alpha^H\}$ defines the basis for $\mathcal{X}(N_\mu/G_\mu)$ we were looking for.

We denote by $(\bar{\mathcal{A}}_\beta^\alpha)$ the matrix inverse to $(\mathcal{A}_\beta^\alpha)$ and set $\Phi^A=\iota^A+\Upsilon_\alpha^A\iota^\alpha$, $\Phi_m^A=\iota^A+\Upsilon_\alpha^A\iota^\alpha+\Upsilon_i^A v^i$, and $\Psi^\alpha=\bar{\mathcal{A}}_\beta^\alpha\iota^\beta$. Then, Γ takes the form

$$\Gamma=\Phi^A\tilde{E}_A^C+\Psi^\alpha\hat{E}_\alpha^H+v^i\bar{X}_i^C+(\Gamma^i\circ\iota)\bar{X}_i^V,$$

$$=\Phi_m^A\tilde{E}_A^C+\Psi^\alpha\hat{E}_\alpha^H+v^i\bar{X}_i^H+(\Gamma^i\circ\iota)\bar{X}_i^V,$$

where, as before, the first term is the vertical part of Γ with respect to the vertical lift connection Ω^{N_μ} (in the first place) and the mechanical connection Ω^m (in the second). Obviously, v^i and $(\Gamma^i\circ\iota)$ are G_μ-invariant functions. To see that Ψ^α is also G_μ-invariant, recall that $C_{AB}^\alpha=0$ and $\tilde{E}_A^C(\iota^\beta)=C_{A\gamma}^\beta\iota^\gamma$ and observe that $\tilde{E}_A^C(\bar{\mathcal{A}}_\delta^\beta)\mathcal{A}_\beta^\gamma=-\bar{\mathcal{A}}_\delta^\beta\tilde{E}_A^C(\mathcal{A}_\beta^\gamma)=-\bar{\mathcal{A}}_\delta^\beta\mathcal{A}_\beta^\alpha C_{A\alpha}^\gamma=-C_{A\delta}^\gamma$. Therefore,

$$\tilde{E}_A^C(\Psi^\alpha)=\tilde{E}_A^C(\bar{\mathcal{A}}_\beta^\alpha)\iota^\beta+\bar{\mathcal{A}}_\beta^\alpha\tilde{E}_A^C(\iota^\beta)=-\bar{\mathcal{A}}_\delta^\alpha C_{A\beta}^\delta\iota^\beta+\bar{\mathcal{A}}_\beta^\alpha C_{A\gamma}^\beta\iota^\gamma=0.$$

We conclude that v^i, $(\Gamma^i\circ\iota)$, and Ψ^α can all be regarded as functions on N_μ/G_μ. The horizontal part of Γ, for both connections, can thus be interpreted as the horizontal lift of the reduced vector field

$$\check{\Gamma} = \Psi^\alpha \check{E}^H_\alpha + v^i \check{X}^C_i + (\Gamma^i \circ \iota)\check{X}^V_i$$

on N_μ / G_μ.

For completeness, we point out that it follows from the relations $\tilde{E}^C_A(\Upsilon^B_\alpha) = C^B_{AC}\Upsilon^C_\alpha - C^\beta_{A\alpha}\Upsilon^B_\beta - C^B_{A\alpha}$ and $\tilde{E}^C_A(\iota^B) = C^B_{AC}\iota^C + C^B_{A\gamma}\iota^\gamma$ that the coefficients Φ^A and Φ^A_m satisfy

$$\tilde{E}^C_A(\Phi^B) = C^B_{AC}\Phi^C, \quad \tilde{E}^C_A(\Phi^B_m) = C^B_{AC}\Phi^C_m.$$

This shows that they can be interpreted as the coefficients of $\mathfrak{g}_\mu$-valued functions Φ and Φ_m on N_μ satisfying $\Phi \circ \psi^{N_\mu}_g = \mathrm{ad}_g \Phi$ for $g \in G_\mu$ (and similarly for Φ_m), where ψ^{N_μ} denotes the G_μ-action on N_μ (see Ref. 3).

We now give a coordinate expression for the reduced vector field. From here on, we will use coordinates $(\theta^a) = (\theta^A, \theta^\alpha)$ such that the fibers of $G \to G/G_\mu$ are given by $\theta^\alpha = \text{const}$. With this assumption, there are functions K^a_b on M such that

$$\tilde{E}_A = K^B_A \frac{\partial}{\partial \theta^B}, \quad \tilde{E}_\alpha = K^B_\alpha \frac{\partial}{\partial \theta^B} + K^\beta_\alpha \frac{\partial}{\partial \theta^\beta}.$$

We also introduce the functions Λ^b_i for which

$$X_i = \frac{\partial}{\partial x^i} - \Lambda^b_i \frac{\partial}{\partial \theta^b},$$

as before.

By interpreting N_μ / G_μ locally as $M/G_\mu \times_{M/G} T(M/G)$, we see that a point of N_μ / G_μ has coordinates $(x^i, \theta^\alpha, v^i)$. Because of their G_μ-invariance, the functions $\Gamma^i \circ \iota$ and Ψ^α are independent of the variables θ^A. Let π^{N_μ} be the projection $N_\mu \to N_\mu / G_\mu$; then for any invariant function F on N_μ, there is a function f on N_μ / G_μ such that $F = f \circ \pi^{N_\mu}$. Then, for all invariant vector fields X on N_μ and their reductions $\check{X}$ to vector fields on N_μ / G_μ, we have

$$X(F) = \check{X}(f) \circ \pi^{N_\mu}.$$

We will apply this property to the vector fields $\hat{E}^H_\alpha$, $\bar{X}^C_i$, and $\bar{X}^V_i$, and the invariant functions x^i, v^i, and θ^α. Keeping in mind that for any vector field Z, function f, and 1-form θ on M,

$$Z^C(\tau^* f) = \tau^* Z(f), \quad Z^V(\tau^* f) = 0, \quad Z^C(\vec{\theta}) = \vec{\mathcal{L}_Z \theta}, \quad Z^V(\vec{\theta}) = \tau^* \theta(Z),$$

where $\vec{\theta}$ stands for the fiber-linear function on TM defined by the 1-form θ, and τ is the tangent projection $TM \to M$, we find that

$$\bar{X}^C_i(x^j) = \delta^j_i, \quad \bar{X}^C_i(\theta^\beta) = -\Lambda^\beta_i, \quad \bar{X}^C_i(v^j) = 0,$$

$$\bar{X}^V_i(x^j) = 0, \quad \bar{X}^V_i(\theta^\beta) = 0, \quad \bar{X}^V_i(v^j) = \delta^j_i,$$

$$\hat{E}^H_\alpha(x^j) = 0, \quad \hat{E}^H_\alpha(\theta^\beta) = \mathcal{A}^\gamma_\alpha K^\beta_\gamma, \quad \hat{E}^H_\alpha(v^j) = 0,$$

from which it follows immediately that

$$\check{X}^C_i = \frac{\partial}{\partial x^i} - \Lambda^\alpha_i \frac{\partial}{\partial \theta^\alpha}, \quad \check{X}^V_i = \frac{\partial}{\partial v^i}, \quad \check{E}^H_\alpha = \mathcal{A}^\gamma_\alpha K^\beta_\gamma \frac{\partial}{\partial \theta^\beta}$$

and

$$\check{\Gamma} = (\iota^\beta K^\alpha_\beta - v^i \Lambda^\alpha_i)\frac{\partial}{\partial \theta^\alpha} + v^i \frac{\partial}{\partial x^i} + (\Gamma^i \circ \iota)\frac{\partial}{\partial v^i}.$$

The equations that determine the integral curves $\check{v}(t)=(x^i(t),\theta^\alpha(t),v^i(t))$ of the reduced vector field $\check{\Gamma}$ are therefore the coupled set

$$\ddot{x}^i = \Gamma^i \circ \iota,$$

$$\dot{\theta}^\alpha = \iota^\beta K^\alpha_\beta - v^i \Lambda^\alpha_i.$$

One can easily convince oneself that the right-hand side of the equation for $\dot{\theta}^\alpha$ is indeed independent of the variables θ^A: by considering the coefficients of $\partial/\partial\theta^\alpha$ in $[\widetilde{E}_A,\widetilde{E}_\beta]=-C^c_{A\beta}\widetilde{E}_c$, we find that $\widetilde{E}_A(K^\alpha_\beta)=-C^\gamma_{A\beta}K^\alpha_\gamma$. Since also $\widetilde{E}^C_A(\iota^\beta)=C^\beta_{A\gamma}\iota^\gamma$ and $\widetilde{E}_A(\Lambda^\beta_i)=0$, it follows easily that $\widetilde{E}^C_A(\iota^\beta K^\alpha_\beta - v^i\Lambda^\alpha_i)=0$, as claimed.

The functions $\Gamma^i\circ\iota$ on the right-hand side of the equation for $\ddot{x}^i$ can be determined from the generalized Routh equations of Sec. III,

$$\Gamma(\bar{X}^V_i(\mathcal{R}^\mu)) - \bar{X}^C_i(\mathcal{R}^\mu) = -\mu_a R^a_{ij} v^j.$$

Since $\mathcal{R}^\mu=\mathcal{R}\circ\iota$ is G_μ-invariant, so also are $\bar{X}^V_i(\mathcal{R}^\mu)$ and $\bar{X}^C_i(\mathcal{R}^\mu)$. Recall that R^a_{ij} are functions on M, determined by $[X_i,X_j]=R^a_{ij}\widetilde{E}_a$. Thus, since $[\widetilde{E}_a,X_i]=0$,

$$(\widetilde{E}_a(R^b_{ij}) - R^c_{ij}C^b_{ac})\widetilde{E}_b = 0.$$

However, $C^b_{Ac}\mu_b=0$, and so

$$\widetilde{E}^C_A(R^b_{ij}\mu_b) = \widetilde{E}_A(R^b_{ij})\mu_b = R^c_{ij}C^b_{Ac}\mu_b = 0.$$

It follows that the term $\mu_a R^a_{ij} v^j$ is G_μ-invariant. The generalized Routh equations therefore pass to the quotient N_μ/G_μ and take the reduced form

$$\check{\Gamma}(\check{X}^V_i(\mathcal{R}^\mu)) - \check{X}^C_i(\mathcal{R}^\mu) = -\mu_a R^a_{ij} v^j.$$

Following Ref. 10, we will call these reduced equations the Lagrange–Routh equations. Under the regularity assumptions we have adopted throughout, the function-valued matrix $(\bar{X}^V_i\bar{X}^V_j(\mathcal{R}^\mu))$ is nonsingular and the coefficients $\Gamma^i\circ\iota$, now interpreted as functions on N_μ/G_μ, can be determined from the Lagrange–Routh equations. In the current coordinate system, the equations become

$$\frac{d}{dt}\left(\frac{\partial\mathcal{R}^\mu}{\partial v^i}\right) - \frac{\partial\mathcal{R}^\mu}{\partial x^i} = -\mu_a R^a_{ij} v^j - \Lambda^\alpha_i \frac{\partial\mathcal{R}^\mu}{\partial\theta^\alpha}.$$

Given a reduced solution $\check{v}(t)=(x^i(t),\theta^\alpha(t),v^i(t))\in N_\mu/G_\mu$, we can apply the method of reconstruction using either one of the connections Ω^m and Ω^{N_μ} to recover a complete solution $v(t)=(x^i(t),\theta^A(t),\theta^\alpha(t),v^i(t))\in N_\mu$ of the Lagrangian system. The examples discussed in Sec. VIII will make it clear how this method works in practice.

VII. SIMPLE MECHANICAL SYSTEMS

In this section, we reconcile our results with those for the case of a simple mechanical system to be found elsewhere in the literature.

A simple mechanical system is one whose Lagrangian is of the form $L=T-V$, where T is a kinetic energy function, defined by a Riemannian metric g on M, and V is a function on M, the potential energy. The symmetry group G consists of those isometries of g which leave V invariant. We define a connection on $M\to M/G$ by taking for horizontal subspaces the orthogonal complements to the tangent spaces to the fibers; it is this connection that is usually called the mechanical connection (for a simple mechanical system). We write $g_{ab}=g(\tilde{E}_a,\tilde{E}_b)$, $g_{ij}=g(X_i,X_j)$; by assumption, $g(\tilde{E}_a,X_i)=0$. Then, in terms of quasivelocities,

$$L(m,v) = \tfrac{1}{2}(g_{ij}(m)v^i v^j + g_{ab}(m)v^a v^b) - V(m),$$

and $\tilde{E}_a^V\tilde{E}_b^V(L)=g_{ab}$, etc., so the notation is consistent with what has gone before. Note that since we assume that g is Riemannian and therefore positive definite, it is automatic that L is regular and that the matrices $(g_{ab}(m))$ and $(g_{ij}(m))$ are both nonsingular for all m; in particular, we do not need to make the separate assumption that (g_{ab}) is nonsingular. Considered as defining a map $M \to \mathfrak{g}^*\odot\mathfrak{g}^*$, (g_{ab}) is called the locked inertia tensor. The isometry condition gives

$$\tilde{E}_a(g_{bc}) + C_{ab}^d g_{cd} + C_{ac}^d g_{bd} = 0, \quad \tilde{E}_a(g_{ij}) = 0.$$

The first of these is the differential version of the equivariance property of the locked inertia tensor with respect to the action of G on M and the coadjoint action of G on $\mathfrak{g}^*\odot\mathfrak{g}^*$. The second shows that g_{ij} may be considered as a function on M/G.

The momentum is given simply by $p_a(m,v)=g_{ab}(m)v^b$. On any level set N_μ, where $p_a=\mu_a$, we can solve explicitly for the v^a to obtain $v^a=g^{ab}\mu_b$.

The Routhian is given by

$$\mathcal{R} = L - p_a v^a = \tfrac{1}{2}g_{ij}v^i v^j - \tfrac{1}{2}g_{ab}v^a v^b - V,$$

and on restriction to N_μ we obtain

$$\mathcal{R}^\mu = \tfrac{1}{2}g_{ij}v^i v^j - \left(V + \tfrac{1}{2}g^{ab}\mu_a\mu_b\right).$$

The quantity $V+\frac{1}{2}g^{ab}\mu_a\mu_b$ is the so-called amended potential[15] and the term $C_\mu=\frac{1}{2}g^{ab}\mu_a\mu_b$ is called the "amendment" in Ref. 10. Both functions on M are G_μ-invariant: one easily verifies that $\tilde{E}_a(C_\mu)=g^{bc}C_{ab}^d\mu_c\mu_d$, so, in particular, for $a=A$ we get $\tilde{E}_A(C_\mu)=0$.

Note that by the choice of connection, $B_i^a=0$; we have $\hat{X}_i^C=\bar{X}_i^C$, and the generalized Routh equations are

$$\Gamma_0(\bar{X}_i^V(\mathcal{R}^\mu)) - \bar{X}_i^C(\mathcal{R}^\mu) = -\mu_a R_{ij}^a v^j.$$

This equation is the analog in our framework of the one in Corollary III.8 of Ref. 10. We have shown in the previous section that it reduces to the Lagrange–Routh equations

$$\check{\Gamma}(\check{X}_i^V(\mathcal{R}^\mu)) - \check{X}_i^C(\mathcal{R}^\mu) = -\mu_a R_{ij}^a v^j,$$

which for consistency should be supplemented by the equation that determines the variables θ^α. As we pointed out earlier, the latter is actually just the expression for genuine velocity components $\dot{\theta}^\alpha$ in terms of quasivelocities, supplemented by the constraint $v^\alpha=\iota^\alpha$ which for a simple mechanical system takes the form $\iota^\alpha=g^{\alpha a}\mu_a$.

We can split the reduced Routhian $\mathcal{R}^\mu$ in a Lagrangian part $\mathfrak{L}=\frac{1}{2}g_{ij}v^i v^j-V$ and the reduced amendment $\mathfrak{C}_\mu$. Since the quasivelocities v^a do not appear in the expression of $\mathfrak{L}$, it can formally be interpreted as a function on $T(Q/G)$. The reduced amendment is a function on Q/G_μ. We can now rewrite the Lagrange–Routh equations as

$$\check{\Gamma}(\check{X}_i^V(\mathfrak{L})) - \check{X}_i^C(\mathfrak{L}) = -\mu_a R_{ij}^a v^j + \check{X}_i(\mathfrak{C}_\mu).$$

In coordinates,

$$\frac{d}{dt}\left(\frac{\partial \mathfrak{L}}{\partial v^i}\right) - \frac{\partial \mathfrak{L}}{\partial x^i} = -\mu_a R^a_{ij} v^j + \left(\frac{\partial}{\partial x^i} - \Lambda^\alpha_i \frac{\partial}{\partial \theta_\alpha}\right)(\mathfrak{C}_\mu).$$

This equation is only one out of two equations that appear in Theorem III.14 in Ref. 10, the theorem that states the reduced equations obtained by following a variational approach to Routh's procedure. We leave it to the reader to verify that the second equation, in its form (III.37), is, in fact,

$$v^\alpha C^a_{\beta\alpha}\mu_a = g^{\alpha b}\mu_b C^a_{\beta\alpha}\mu_a.$$

Since $v^\alpha = g^{\alpha b}\mu_b$, this is obviously an identity from the current point of view; it certainly cannot be used to determine $\dot{\theta}^\alpha$ in terms of the other variables, and without this information the equations are incomplete. In this respect, therefore, our reduction results are an improvement on those in Ref. 10.

Let us now check, in the case where the configuration space M is of the form $S\times G$, for an Abelian symmetry group ($C^c_{ab}=0$), and a Lagrangian of the form

$$L(x,\theta,\dot{x},\dot{\theta}) = \tfrac{1}{2}k_{ij}(x)\dot{x}^i\dot{x}^j + k_{ia}(x)\dot{x}^i\dot{\theta}^a + \tfrac{1}{2}k_{ab}(x)\dot{\theta}^a\dot{\theta}^b - V(x),$$

that the reduced equations above coincide with those in the Introduction. We set

$$\tilde{E}_a = K^b_a \frac{\partial}{\partial \theta^b},$$

where the K^a_b are independent of the θ^a since we are dealing with the Abelian case. In general, horizontal vector fields take the form

$$X_i = \frac{\partial}{\partial x^i} - \Lambda^a_i \frac{\partial}{\partial \theta^a}.$$

The quasivelocities adapted to the connection are therefore given, as before, by $v^i = \dot{x}^i$ and $K^a_b v^b = \dot{\theta}^a + \Lambda^a_i \dot{x}^i$.

Given that in this case $\tilde{E}_a(g_{bc})=0$ and $\tilde{E}_a(g_{ij})=0$, all coefficients of the metric can be interpreted as functions on $M/G=S$, and they depend only on the variables x^i. The use of the mechanical connection entails that $g_{ai}=0$. When expressed in terms of the coordinates $(\dot{x}^i, \dot{\theta}^a)$, this property fixes the connection coefficients to be of the form $\Lambda^a_i = k^{ab}k_{ib}$ and the remaining coefficients of the metric to be $g_{ab}=k_{cd}K^c_aK^d_b$ and $g_{ij}=k_{ij}-k^{ab}k_{ia}k_{jb}$. The expression for $\mathcal{R}^\mu$ given in the Introduction now easily follows. Since $\mathcal{R}^\mu$ is a function only of x^i and $v^i=\dot{x}^i$, we get

$$\bar{X}^V_i(\mathcal{R}^\mu) = \frac{\partial \mathcal{R}^\mu}{\partial \dot{x}^i} \quad \text{and } \bar{X}^C_i(\mathcal{R}^\mu) = \frac{\partial \mathcal{R}^\mu}{\partial x^i}.$$

Moreover,

$$[X_i, X_j] = B^a_{ij}\frac{\partial}{\partial \theta^a} = R^a_{ij}\tilde{E}_a,$$

and likewise $\mu_a = \tilde{E}^V_a(L) = K^b_a\pi_b$. So, $\mu_a R^a_{ij} = \pi_a B^a_{ij}$ and the equation from the Introduction follows.

We now return to the general case (of a simple mechanical system) and consider the reconstruction process.

We continue to use the mechanical connection on M. Since, in the basis that is adapted to this connection, $g_{ia}=0$, and therefore $\bar{X}^H_i = \bar{X}^C_i$, the two connections Ω^m and Ω^{N_μ} coincide. We denote the common connection on N_μ by Ω.

Let $\check{v}(t)$ be a curve in N_μ/G_μ which is an integral curve of $\check{\Gamma}$, and $\check{v}^H$ a horizontal lift of $\check{v}$ to N_μ (horizontal with respect to Ω). The reconstruction equation is

$$\widetilde{\vartheta(\dot{g})}^C = \Omega(\Gamma \circ \check{v}^H)$$

(where ϑ here is the Maurer–Cartan form of G_μ); this is (at each point on the curve $\check{v}$) an equation between vertical vectors on N^μ but can and should be thought of as an equation on $\mathfrak{g}_\mu$. It determines a curve $g(t)$ in G_μ such that

$$t \mapsto \psi^{N_\mu}_{g(t)} \check{v}^H(t)$$

is an integral curve of Γ in N_μ; again, ψ^{N_μ} is the action of G_μ on N_μ. So far, this works for an arbitrary Lagrangian.

Now, as we showed earlier in general, the vertical part of Γ with respect to the vertical lift connection Ω^{N_μ} is $(\iota^A + \Upsilon^A_\alpha \iota^\alpha)\widetilde{E}^C_A$. Thus, in the case at hand,

$$\Omega(\Gamma) = (\iota^A + \Upsilon^A_\alpha \iota^\alpha)\widetilde{E}^C_A = (\iota^A + G^{AB} g_{B\alpha}\iota^\alpha)\widetilde{E}^C_A = (g^{Aa}\mu_a + G^{AB} g_{B\alpha} g^{\alpha a}\mu_a)\widetilde{E}^C_A = (g^{Aa}\mu_a + G^{AB}(\delta^a_B - g_{BC}g^{Ca})\mu_a)\widetilde{E}^C_A = (g^{Aa}\mu_a + G^{AB}\mu_B - g^{Aa}\mu_a)\widetilde{E}^C_A = G^{AB}\mu_B \widetilde{E}^C_A.$$

The first point to note is that the coefficient $G^{AB}\mu_B$ appearing on the right-hand side of the final equation above is a function on M, so that in the right-hand side of the reconstruction equation the argument $\check{v}^H$ can be replaced by its projection into M, which is $\tau \circ \iota \circ \check{v}^H$.

Next, we interpret $G^{AB}\mu_B$ in terms of the locked inertia tensor. Recall that the locked inertia tensor at $m \in M$ has components $g_{ab}(m)$. As is the usual practice, we consider the locked inertia tensor as a nonsingular symmetric linear map $I(m): \mathfrak{g} \to \mathfrak{g}^*$. Now, let j be the injection $\mathfrak{g}_\mu \to \mathfrak{g}$: then, μ_B are the components of $j^*\mu \in \mathfrak{g}^*_\mu$ and $g_{AB}(m)$ are the components of the map $I_\mu(m) = j^* \circ I(m) \circ j$. Then,

$$G^{AB}(m)\mu_B E_A = I^{-1}_\mu(m)(j^*\mu),$$

a point of $\mathfrak{g}_\mu$. So, finally, the reconstruction equation may be written as

$$\vartheta(\dot{g}(t)) = I^{-1}_\mu(c(t))(j^*\mu), \quad c = \tau \circ \iota \circ \check{v}^H.$$

This is an equation between curves in $\mathfrak{g}_\mu$.

We will now show that this reconstruction equation above is a particular and simple case of one of the reconstruction equations appearing in Ref. 10.

To do so, we must introduce yet another connection, used in Ref. 10 and called there the mechanical connection for the G^μ-action. This is a connection on the principal fiber bundle $M \to M/G_\mu$, i.e., a G_μ-invariant splitting of the short exact sequence

$$0 \to M \times \mathfrak{g}_\mu \to TM \to M \times_{M/G_\mu} T(M/G_\mu) \to 0.$$

If, as before, $\{E_A, E_\alpha\}$ is a basis of $\mathfrak{g}$ for which $\{E_A\}$ is a basis for $\mathfrak{g}_\mu$, then the vector fields X_i together with the vector fields $\widetilde{E}_\alpha - G^{AB} g_{A\alpha}\widetilde{E}_B$ form a basis for the set of vector fields which are horizontal with respect to the mechanical connection for the G^μ-action. We denote the latter by ω^μ. Now, ω^μ and Ω are related somewhat as a connection and its vertical lift: in fact (for their projections onto $\mathfrak{g}_\mu$),

$$\Omega(Z_v) = \omega^\mu(T(\tau \circ \iota)Z_v), \quad Z_v \in TN_\mu.$$

We note in passing that since $T(\tau \circ \iota)\Gamma(v) = v$ for any $v \in N_\mu$, we can write the reconstruction equation as

$$\vartheta(\widetilde{\dot{g}(t)}) = \omega^\mu(\check{v}^H).$$

The reconstruction equation in Ref. 10 that we are aiming for is the third of the four, Eq. (IV.6). It seems the one most relevant to our approach because, as Marsden *et al.* said, in it they "take the dynamics into account," and this has been our purpose throughout. Now, Eq. (IV.6) of Ref. 10

differs from our reconstruction equation (expressed in terms of I_μ) by having an additional term on the right-hand side involving the mechanical connection for the G^μ-action ω^μ. This arises because the authors start with a more general class of curves on M than we do. In order to show that our equation agrees with theirs, we first show that the curve $c=\tau\circ\iota\circ\check{v}^H$ in M is ω^μ-horizontal; the extra term in their equation is therefore zero in our case. By evaluating ω^μ on the tangent to c and using the relation between ω^μ and Ω, we have

$$\omega^\mu(\dot{c}) = \omega^\mu(T(\tau\circ\iota)\check{v}^H) = \Omega(\check{v}^H) = 0$$

because $\check{v}^H$ is Ω-horizontal. So, our reconstruction equation formally agrees with Eq. (IV.6) of Marsden *et al.*, when we take the starting curve on M to be c: it is the particular case of that equation in which the curve on M is horizontal with respect to the G_μ mechanical connection.

To finish the story, we must also take into account the fact that Eq. (IV.6) of Ref. 10 is presented as an equation for the reconstruction of a base integral curve of Γ, with momentum μ, from another suitable curve on M, whereas our reconstruction equation gives an integral curve of Γ on N^μ. However, there is no real discrepancy here because Γ is a second-order differential equation field and so knowing its base integral curves is equivalent to knowing its integral curves. Let us spell this out in detail. We know that if $t\mapsto g(t)$ is a solution of our reconstruction equation, then

$$t \mapsto \psi^{N_\mu}_{g(t)}\check{v}^H(t)$$

is an integral curve of Γ in N_μ. The corresponding base integral curve is

$$t \mapsto \tau(\iota(\psi^{N_\mu}_{g(t)}\check{v}^H(t))).$$

However,

$$\tau\circ\iota\circ\psi^{N_\mu}_g = \tau\circ\psi^{TM}_g\circ\iota = \psi^M_g\circ\tau\circ\iota,$$

so the curve $t\mapsto\psi^M_{g(t)}c(t)$ is a base integral curve of Γ. Thus, the same curve in G_μ determines an integral curve of Γ in N_μ (by its action on $\check{v}^H$) and the corresponding base integral curve (by its action on $c=\tau\circ\iota\circ\check{v}^H$, the projection of $\check{v}^H$ to M).

VIII. ILLUSTRATIVE EXAMPLES

We give two examples. In the first, we derive Wong's equations using our methods. This example is intended to illustrate the Routhian approach in a case of some physical interest; however, we do not pursue the calculations as far as the consideration of the isotropy algebra and reconstruction. These matters are illustrated in the second example, which is more specific and more detailed, if somewhat more artificial.

A. Wong's equations

We discuss the generalized Routh equations for the geodesic field of a Riemannian manifold on which a group G acts freely and properly to the left as isometries, and where the vertical part of the metric (that is, its restriction to the fibers of $\pi^M: M\to M/G$) comes from a bi-invariant metric on G. The reduced equations in such a case are known as Wong's equations.[2,12]

This is of course an example of a simple mechanical system, with $V=0$; we therefore adopt the notation of Sec. VII, and we will use the mechanical connection. In order to utilize conveniently the assumption about the vertical part of the metric g, we will need symbols for the components of g with respect to the invariant vector fields $\hat{E}_a$ introduced in Sec. II; we write

$$h_{ab} = g(\hat{E}_a, \hat{E}_b) = \mathcal{A}_a^c \mathcal{A}_b^d g_{cd}.$$

Since both h_{ab} and g_{ij} are G-invariant functions, they pass to the quotient. In particular, g_{ij} are the components with respect to the coordinate fields of a metric on M/G, the reduced metric; we denote by Γ_{ij}^k its Christoffel symbols.

The further assumption about the vertical part of the metric has the following implications. It means in the first place that $\mathcal{L}_{\hat{E}_c} g(\hat{E}_a, \hat{E}_b) = 0$ [as well as $\mathcal{L}_{\tilde{E}_c} g(\hat{E}_a, \hat{E}_b) = 0$]. Taking into account the bracket relations $[\hat{E}_a, \hat{E}_b] = C_{ab}^c \hat{E}_c$, we find that h_{ab} must satisfy $h_{ad} C_{bc}^d + h_{bd} C_{ac}^d = 0$. It is implicit in our choice of an invariant basis that we are working in a local trivialization of $M \to M/G$. Then, h_{ab} are functions on the G factor, so must be independent of the coordinates x^i on M/G, which is to say that they must be constants. Moreover, $\tilde{E}_a$, $\hat{E}_a$, and $\mathcal{A}_a^b$ are all objects defined on the G factor, so are independent of the x^i. We may write

$$X_i = \frac{\partial}{\partial x^i} - \gamma_i^a \hat{E}_a$$

for some coefficients γ_i^a which are clearly G-invariant; moreover, $[X_i, \hat{E}_a] = \gamma_i^c C_{ac}^b \hat{E}_b$. We set $\gamma_i^c C_{ac}^b = \gamma_{ia}^b$; then, $h_{ac} \gamma_{ib}^c + h_{bc} \gamma_{ib}^c = 0$.

We are interested in the geodesic field of the Riemannian metric g. The geodesic equations may be derived from the Lagrangian

$$L = \tfrac{1}{2} g_{\alpha\beta} u^\alpha u^\beta = \tfrac{1}{2} g_{ij} v^i v^j + \tfrac{1}{2} g_{ab} v^a v^b = \tfrac{1}{2} g_{ij} v^i v^j + \tfrac{1}{2} h_{ab} w^a w^b,$$

where w^a are quasivelocities relative to $\hat{E}_a$; we have $\mathcal{A}_b^a w^b = v^a$. The momentum is given by $p_a = g_{ab} v^b = \bar{\mathcal{A}}_a^c h_{bc} w^c$, where $(\bar{\mathcal{A}}_a^b)$ is the matrix inverse to $(\mathcal{A}_a^b)$. The Routhian is

$$\mathcal{R} = \tfrac{1}{2} g_{ij} v^i v^j - \tfrac{1}{2} g^{ab} p_a p_b.$$

It is easy to see that $\bar{X}_i^V(\mathcal{R}) = g_{ij} v^j$. The calculation of $\bar{X}_i^C(\mathcal{R})$ reduces to the calculation of $X_i(g_{ij})$ and $X_i(g^{ab})$. The first is straightforward. For the second, we note that $g_{ab} = \bar{\mathcal{A}}_a^c \bar{\mathcal{A}}_b^d h_{cd}$; since the right-hand side is independent of x^i, so is g_{ab}, and so equally is g^{ab}. It follows that

$$\bar{X}_i^C(\mathcal{R}) = \frac{1}{2} \frac{\partial g_{jk}}{\partial x^i} v^j v^k - \frac{1}{2} \gamma_i^c \hat{E}_c(g^{ab}) p_a p_b.$$

Now, $\hat{E}_c(g^{ab}) = -\mathcal{A}_c^d (g^{ae} C_{de}^b + g^{be} C_{de}^a)$, from Killing's equations. Using the relation between g_{ab} and h_{ab}, and the fact that ad is a Lie algebra homomorphism, we find that

$$\hat{E}_c(g^{ab}) = -A_d^a A_e^b (h^{df} C_{cf}^e + h^{ef} C_{cf}^d).$$

The expression in the brackets vanishes, as follows easily from the properties of h_{ab}. Thus, the generalized Routh equation is

$$\frac{d}{dt}(g_{ij} v^j) - \frac{1}{2} \frac{\partial g_{jk}}{\partial x^i} v^j v^k = g_{ij}(\dot{v}^j + \Gamma_{kl}^j v^k v^l) = -\mu_a R_{ij}^a v^j.$$

However, $\mu_a = g_{ab} v^b = \bar{\mathcal{A}}_a^c h_{bc} w^b$; so if we set $K_{ij}^a = \bar{\mathcal{A}}_b^a R_{ij}^b$, then $\mu_a R_{ij}^a = h_{bc} K_{ij}^c w^b$. The generalized Routh equation is therefore equivalent to

$$\ddot{x}^i + \Gamma_{jk}^i \dot{x}^j \dot{x}^k = g^{im} h_{bc} K_{lm}^c \dot{x}^l w^b.$$

We also need an equation for w^a: this comes from the constancy of μ_a, which we may write as

$$h_{bc}\frac{d}{dt}(\bar{\mathcal{A}}_a^c w^b) = 0.$$

If we are to understand this equation in the present context, we evidently need to calculate $\dot{\mathcal{A}}_a^b$. Now,

$$\dot{\mathcal{A}}_a^b = v^i X_i(\mathcal{A}_a^b) + v^c \widetilde{E}_c(\mathcal{A}_a^b) = v^i \gamma_{ia}^c \mathcal{A}_c^b + v^c C_{cd}^b \mathcal{A}_a^d.$$

It follows that

$$h_{bc}\frac{d}{dt}(\bar{\mathcal{A}}_a^c) = - h_{bc}\bar{\mathcal{A}}_a^d \bar{\mathcal{A}}_e^c \dot{\mathcal{A}}_d^e = - h_{bc}\bar{\mathcal{A}}_a^d \bar{\mathcal{A}}_e^c (v^i \gamma_{id}^f \mathcal{A}_f^e + v^f C_{fg}^e \mathcal{A}_d^g) = - h_{bc}\bar{\mathcal{A}}_a^d (v^i \gamma_{id}^c + w^e C_{ed}^c),$$

where in the last step we have again used the fact that ad is a Lie algebra homomorphism. Now, from the skew-symmetry properties of h_{ab}, we obtain

$$h_{bc}\frac{d}{dt}(\bar{\mathcal{A}}_a^c) = h_{cd}\bar{\mathcal{A}}_a^d (v^i \gamma_{ib}^c + w^e C_{eb}^c),$$

and therefore

$$h_{bc}\frac{d}{dt}(\bar{\mathcal{A}}_a^c w^b) = h_{cd}\bar{\mathcal{A}}_a^d (\dot{w}^c + \gamma_{ib}^c v^i w^b).$$

The generalized Routh equation and the constancy of momentum together amount to the mixed first- and second-order equations

$$\ddot{x}^i + \Gamma_{jk}^i \dot{x}^j \dot{x}^k = g^{im} h_{bc} K_{lm}^c \dot{x}^l w^b,$$

$$\dot{w}^a + \gamma_{jb}^a \dot{x}^j w^b = 0.$$

These are Wong's equations as they are usually expressed.

B. A Lagrangian with SE (2) as symmetry group

We now consider the Lagrangian (of simple mechanical type)

$$L = \tfrac{1}{2}\dot{x}^2 + \tfrac{1}{2}\dot{y}^2 + \tfrac{1}{2}\dot{z}^2 + \tfrac{1}{2}\dot{\theta}^2 + A((\sin\theta)\dot{z} + (\cos\theta)\dot{y})\dot{\theta}.$$

The system is regular if $A^2 \neq 1$. The Euler–Lagrange equations are

$$\ddot{x} = 0, \quad \frac{d}{dt}(\dot{y} + (A\cos\theta)\dot{\theta}) = 0, \quad \frac{d}{dt}(\dot{z} + (A\sin\theta)\dot{\theta}) = 0, \quad \ddot{\theta} + (A\sin\theta)\ddot{z} + (A\cos\theta)\ddot{y} = 0,$$

and the solution with (for convenience) $\theta_0 = 0$ is

$$(x(t), y(t), z(t), \theta(t)) = (\dot{x}_0 t + x_0, - A\sin(\dot{\theta}_0 t) + (\dot{y}_0 + A\dot{\theta}_0)t + y_0, A\cos(\dot{\theta}_0 t) + \dot{z}_0 t + z_0 - A, \dot{\theta}_0 t).$$

The system is invariant under the group SE (2), the special Euclidean group of the plane. The configuration manifold is $\mathbf{R}\times \mathrm{SE}(2)$, where x is the coordinate on $\mathbf{R}$. We will use the trivial connection. An element of SE (2) can be represented by the matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta & y \\ \sin\theta & \cos\theta & z \\ 0 & 0 & 1 \end{pmatrix}.$$

The identity of the group is ($y=0$, $z=0$, $\theta=0$) and the multiplication is given by

$$(y_1,z_1,\theta_1)*(y_2,z_2,\theta_2)=(y_2\cos\theta_1-z_2\sin\theta_1+y_1,y_2\sin\theta_1+z_2\cos\theta_1+z_1,\theta_1+\theta_2).$$

The matrices,

$$e_1=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix},\quad e_2=\begin{pmatrix}0&0&0\\0&0&1\\0&0&0\end{pmatrix},\quad e_3=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix},$$

form a basis for the Lie algebra, for which $[e_1,e_2]=0$, $[e_1,e_3]=e_2$, and $[e_2,e_3]=-e_1$. The corresponding basis for the fundamental vector fields is

$$\tilde{e}_1=\frac{\partial}{\partial y},\quad \tilde{e}_2=\frac{\partial}{\partial z},\quad \tilde{e}_3=-z\frac{\partial}{\partial y}+y\frac{\partial}{\partial z}+\frac{\partial}{\partial\theta},$$

and for the invariant vector fields, we get

$$\hat{e}_1=\cos\theta\frac{\partial}{\partial y}+\sin\theta\frac{\partial}{\partial z},\quad \hat{e}_2=-\sin\theta\frac{\partial}{\partial y}+\cos\theta\frac{\partial}{\partial z},\hat{e}_3=\frac{\partial}{\partial\theta}.$$

One can easily verify that the Lagrangian is invariant.

Before we calculate an expression for the level sets $p_a=\mu_a$, we will examine the isotropy algebra $\mathfrak{g}_\mu$ of a generic point $\mu=\mu_1e^1+\mu_2e^2+\mu_3e^3$ in $\mathfrak{g}^*$. The relations that characterize an element $\xi=\xi^1e_1+\xi^2e_2+\xi^3e_3$ of $\mathfrak{g}_\mu$ are

$$\xi^3\mu_2=0,\quad \xi^3\mu_1=0,\quad \xi^1\mu_2-\xi^2\mu_1=0.$$

So, if we suppose that μ_1 and μ_2 do not both vanish, we will take them from now on to be 1 and μ, respectively, then a typical element of $\mathfrak{g}_\mu$ is $\xi=\xi^1(e_1+\mu e_2)$. We will also set $\mu_3=0$ for convenience. Since $\mathfrak{g}_\mu$ is one dimensional, it is of course Abelian.

Before writing down the coordinate version of the reduced equations in the previous sections, we made two assumptions. First, we supposed that a part of the basis of $\mathfrak{g}$ was, in fact, a basis of $\mathfrak{g}_\mu$. So from now on, we will work with a new basis $\{E_1=e_1+\mu e_2,\ E_2=e_2,\ E_3=e_3\}$, with corresponding notations for the fundamental and invariant vector fields. The Lie algebra brackets in this basis are $[E_1,E_2]=0$, $[E_1,E_3]=-\mu E_1+(1+\mu^2)E_2$, and $[E_2,E_3]=-E_1+\mu E_2$. The momentum vector with which we are working takes the form $(1+\mu^2)E^1+\mu E^2$ (with $\mu_3=0$), when written with respect to the new dual basis.

The second assumption is that we use coordinates $(\theta^a)=(\theta^A,\theta^\alpha)$ on G such that the fibers $G\to G/G_\mu$ are given by θ^α=const. Then, fundamental vector fields for the G_μ-action on G are of the form $K_A^B\partial/\partial\theta^B$. The main advantage of this assumption is that in these coordinates, the expressions in the reduced equations became independent of the coordinates θ^A. This assumption is not yet satisfied in our case for the coordinates (y,z,θ). The action of G_μ on G is given by the restriction of the multiplication, i.e., by

$$(y_1)*(y_2,z_2,\theta_2)=(y_2+y_1,z_2,\theta_2).$$

We have only one coordinate on G_μ, say, y'. The fundamental vector fields that correspond to this action should be of the form $K\partial/\partial y'$. However, in the new basis, vectors in $\mathfrak{g}_\mu$ are of the form KE_1, with corresponding fundamental vector fields

$$K\tilde{E}_1=K\left(\frac{\partial}{\partial y}+\mu\frac{\partial}{\partial z}\right).$$

So, we should make a coordinate change $(y,z,\theta)\to(y',z',\theta')$, such that

$$\frac{\partial}{\partial y'} = \frac{\partial}{\partial y} + \mu \frac{\partial}{\partial z}.$$

This can be done by putting

$$y' = y, \quad z' = z - \mu y, \quad \theta' = \theta.$$

We will then have coordinates $(y', z', \theta', x, \dot{x})$ on N_μ and $(z', \theta', x, \dot{x})$ on N_μ/G_μ. To save typing, we will use y and θ for y' and θ', and only make the distinction between z and z'.

The first goal is to solve the reduced equations on N_μ/G_μ. They are of the form

$$\ddot{x}^i = \Gamma^i(x^j, \theta^\alpha, \dot{x}^j),$$

$$\dot{\theta}^\alpha = \iota^\beta K^\alpha_\beta - \dot{x}^i \Lambda^\alpha_i.$$

For this example, there is only one coordinate x on $\mathbf{R}$ [we are using the trivial connection on $\mathrm{SE}(2) \times \mathbf{R} \to \mathbf{R}$], but the coordinates θ^α on $\mathrm{SE}(2)/G_\mu$ are (z', θ). The reduced second-order equation in x above can be derived from the Lagrangian equation in x which is simply

$$\ddot{x} = 0.$$

It is therefore not coupled to the first-order equation in (z', θ), and its solution is $x(t) = \dot{x}_0 t + x_0$. For the other equations, we will work first with the variables (y, z, θ) and only make the change to the new coordinates at the end.

The matrix (K^α_β) in the above expressions is determined by the relation $\widetilde{E}_a = K^b_a \partial / \partial \theta^a$. It is the lower right (2,2) matrix of

$$K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -z & y + \mu z & 1 \end{pmatrix}.$$

With the trivial connection, the equations for the other variables on N_μ/G_μ are therefore of the form

$$\dot{z}' = \iota^2 + (y + \mu z)\iota^3, \quad \dot{\theta} = \iota^3.$$

We can find the functions ι^a by solving the expressions $p_a = \mu_a$ for v^a, with $(\mu_1, \mu_2, \mu_3) \in \mathfrak{g}^*$ of the form $((1+\mu^2), \mu, 0)$. We get

$$1 + \mu^2 = \dot{y} + \mu \dot{z} + (A \cos \theta + A\mu \sin \theta)\dot{\theta},$$

$$\mu = \dot{z} + (A \sin \theta)\dot{\theta},$$

$$0 = (A \cos \theta)\dot{y} + (A \sin \theta)\dot{z} + \dot{\theta} - z(\dot{y} + (A \cos \theta)\dot{\theta}) + y(\dot{z} + (A \sin \theta)\dot{\theta}).$$

At $t=0$, the above equations relate the integration constants and μ. We will set from now on $\dot{y}_0 = 1 - A\dot{\theta}_0$, $\dot{z}_0 = \mu$, and $z_0 = \mu y_0 + A\dot{y}_0 + \dot{\theta}_0$. It is easy to see that the coordinates v^a with respect to the basis $\{\widetilde{E}_1\}$ (with the trivial connection) are given by

$$v^1 = \dot{y} + z\dot{\theta}, \quad v^2 = \dot{z} - \mu \dot{y} - \mu z \dot{\theta} - y\dot{\theta}, \quad v^3 = \dot{\theta}.$$

After substituting this into the equations for the level set, we obtain the expressions $v^a = \iota^a$ as functions of (y, z, θ). After some calculation, the reduced equations become

$$\dot{z}' = \frac{A}{A^2-1}((z-\mu y)(\sin\theta - \mu\cos\theta) - A(1-\mu^2)\sin\theta\cos\theta - \mu A + 2\mu A(\cos\theta)^2),$$

$$\dot{\theta} = \frac{1}{A^2-1}(\mu y - z + A\cos\theta + A\mu\sin\theta).$$

Observe that we can now replace $(z-\mu y)$ everywhere by the new coordinate z', so that indeed the G_μ-coordinate y' does not appear in the reduced equations. One can verify that the solution of the above equations, with the integration constants determined by μ, is

$$(z'(t),\theta(t)) = (A\cos(\dot{\theta}_0 t) + A\mu\sin(\dot{\theta}_0 t) + (1-A^2)\dot{\theta}_0, \dot{\theta}_0 t).$$

We will now use the mechanical connection to reconstruct the G_μ-part $y(t)$ of the solution. The Hessian of the Lagrangian, in the basis $\{X=\partial/\partial x, \widetilde{E}_a\}$, is

$$\begin{pmatrix} 1+\mu^2 & \mu & A\cos\theta + A\mu\sin\theta - z + \mu y & 0 \\ \mu & 1 & A\sin\theta + y & 0 \\ A\cos\theta + A\mu\sin\theta - z + \mu y & A\sin\theta + y & 1-2Az\cos\theta + 2Ay\sin\theta + y^2 + z^2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The determinant of the matrix is $1-A^2$. The vector field $\bar{X}^C = X^C = \partial/\partial x$ is tangent to the level sets and horizontal with respect to the mechanical connection Ω^m.

In general, we regard the Hessian as a tensor field along the tangent bundle projection. A basis of vector fields along τ that lie in the g-complement of $\mathfrak{g}_\mu$ is

$$\left\{\widetilde{E}_2 - \frac{\mu}{1+\mu^2}\widetilde{E}_1, \widetilde{E}_3 - \frac{1}{1+\mu^2}(A\cos\theta + A\mu\sin\theta - z')\widetilde{E}_1, \frac{\partial}{\partial x}\right\}.$$

Notice that they are all basic vector fields along τ (i.e., vector fields on M). The reason is that the Lagrangian is of the simple type. We have seen that in that case, the g-complement of $\mathfrak{g}_\mu$ defines a connection ω^μ on $M \to M/G_\mu$. The connection tensor Ω^m of the mechanical connection is determined by

$$\Omega^m(\widetilde{E}_1^C) = \widetilde{E}_1^C, \quad \Omega^m(\bar{E}_\alpha^H) = 0, \quad \Omega^m(\hat{X}^C) = 0, \quad \Omega^m(\bar{X}^V) = 0,$$

where the vector fields $\{\bar{E}_\alpha^H\}$ that are horizontal with respect to the mechanical connection Ω^m and tangent to the level set are here

$$\bar{E}_2^H = \bar{E}_2^C - \frac{\mu}{1+\mu^2}\widetilde{E}_1^C, \quad \bar{E}_3^H = \bar{E}_3^C - \frac{1}{1+\mu^2}(A\cos\theta + A\mu\sin\theta - z')\widetilde{E}_1^C.$$

The explicit expressions of $\bar{E}_\alpha^C$ are not of direct concern; we only need to know that they are tangent to the level set and that they differ from $\hat{E}_\alpha^C$ in a vertical lift. The vertical part of $\Gamma = \iota^a \bar{E}_a^C + v^i \bar{X}_i^C + \Gamma^i \bar{X}_i^V$ (the restriction of the dynamical vector field to N_μ) is here

$$\Omega^m(\Gamma) = \left(\iota^1 + \frac{\mu}{1+\mu^2}\iota^2 + \frac{1}{1+\mu^2}(A\cos\theta + A\mu\sin\theta - z')\iota^3\right)\widetilde{E}_1^C.$$

Before we can write down the explicit form of the reconstruction equation $g^{-1}\dot{g} = \Omega^m(\Gamma \circ \breve{v}^H)$, we need to find the horizontal lift $\breve{v}^H$ of the reduced solution $\breve{v} = (z', \theta, x, \dot{x})$. It is the curve $(y_m, z', \theta, x, \dot{x})$ in N_μ whose tangent vector is horizontal with respect to the G_μ-mechanical connection. By construction this means that $(d/dt)(\tau \circ \breve{v}^H)$ should be ω^μ-horizontal. If we write in general that $(d/dt)(\tau \circ \breve{v}^H) = v^1\widetilde{E}_1 + v^2\widetilde{E}_2 + v^3\widetilde{E}_3$, then in order for the curve to be horizontal, the v^a must satisfy

$$v^1 = -v^2\frac{\mu}{1+\mu^2} - v^3\frac{1}{1+\mu^2}(A\cos\theta + A\mu\sin\theta - z').$$

By expressing v^a as functions of $\dot{\theta}^a$, we find that the missing y^H is a solution of

$$\dot{y}^H = -A\dot{\theta}_0\cos(\dot{\theta}_0 t),$$

from which $y^H(t) = -A\sin(\dot{\theta}_0 t) + y_0$. Using this y^H in the reconstruction equation gives

$$\dot{y}_1 = \iota^1 + \frac{\mu}{1+\mu^2}\iota^2 + \frac{1}{1+\mu^2}(A\cos\theta + A\mu\sin\theta - z')\iota^3 = 1,$$

once we have evaluated the functions ι^a in terms of $(y^H, z', \theta, x, \dot{x})$. So, the solution through the identity is $y_1(t) = t$. The y-part of the complete solution of the Euler–Lagrange equation is therefore

$$y(t) = y_1(t) + y^H(t) = -A\sin(\dot{\theta}_0 t) + t + y_0,$$

as it should be for the given value of the momentum.

ACKNOWLEDGMENTS

The first author is a Guest Professor and Ghent University: he is grateful to the Department of Mathematical Physics and Astronomy at Ghent for its hospitality.

The second author is currently a research fellow at The University of Michigan through a Marie Curie Fellowship. He is grateful to the Department of Mathematics for its hospitality. He also acknowledges a research grant (Krediet aan Navorsers) from the Fund for Scientific Research-Flanders (FWO-Vlaanderen), where he is an Honorary Postdoctoral Fellow.

[1] Castrillon-Lopez, M. (unpublished).
[2] Cendra, H., Marsden, J. E., and Ratiu, T. S., *Lagrangian Reduction by Stages*, Memoirs of the American Mathematical Society Vol. 152 (AMS, Providence, 2001).
[3] Crampin, M., and Mestdag, T., "Reduction and reconstruction aspects of second-order dynamical systems with symmetry," Preprint, 2006 (unpublished) (available at maphyast.ugent.be).
[4] Crampin, M., and Pirani, F. A. E., *Applicable Differential Geometry*, LMS Lecture Notes Vol. 59 (Cambridge University Press, Cambridge, 1988).
[5] Jalnapurkar, S. M., and Marsden, J. E., "Reduction of Hamilton's variational principle," Dyn. Stab. Syst. **15**, 287 (2000).
[6] Kobayashi, S., and Nomizu, K., *Foundations of Differential Geometry* (Wiley, New York, 1963), Vol. I.
[7] Marsden, J. E., Montgommery, R., and Ratiu, T., *Reduction, Symmetry and Phases in Mechanics*, Memoirs of the American Mathematical Society Vol. 88 (AMS, Providence, 1990).
[8] Marsden, J. E., and Ratiu, T., *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics Vol. 17 (Springer, New York, 1999).
[9] Marsden, J. E., and Scheurle, J., "Lagrangian reduction and the double spherical pendulum," Z. Angew. Math. Phys. **44**, 17 (1993).
[10] Marsden, J. E., Ratiu, T., and Scheurle, J., "Reduction theory and the Lagrange-Routh equations," J. Math. Phys. **41**, 3379 (2000).
[11] Mestdag, T., and Crampin, M., "Invariant Lagrangians, mechanical connections and the Lagrange-Poincaré equations," J. Phys. A: Math. Theor. (in press).
[12] Montgomery, R., "Canonical formulations of a classical particle in a Yang-Mills field and Wong's equations," Lett. Math. Phys. **8**, 59 (1984).
[13] Routh, E. J., *A Treatise on the Stability of a Given State of Motion* (MacMillan, London, 1877); available on google.books.com
[14] Sharpe, R. W., *Differential Geometry*, Graduate Texts in Mathematics Vol. 166 (Springer, New York, 1997).
[15] Simo, J. C., Lewis, D., and Marsden, J. E., "Stability of relative equilibria. Part I: The reduced energy-momentum method," Arch. Ration. Mech. Anal. **115**, 15 (1991).
[16] Yano, K., and Ishihara, S., *Tangent and Cotangent Bundles* (Dekker, New York, 1973).